INFANT PROSODIC EXPRESSIONS IN MOTHER - INFANT COMMUNICATION

by

CHRISTINA PAPAELIOU

DOCTOR OF PHILOSOPHY

THE UNIVERSITY OF EDINBURGH

1998

DECLARATION

I, hereby, declare that the present thesis has been composed by myself, and is my own work.

TABLE OF CONTENTS

ACKNOWLEDGEMENTS

I am indebted to Professor Colwyn Trevarthen for his invaluable advice on my thesis, his patience, understanding and moral support throughout my entire work. I am grateful to Dr. Morag Donaldson for her great help, useful comments and encouragement. I would like to offer my sincere thanks to my university teachers in Greece, Professor Giannis Kugiumutzakis and Dr. Korina Koufakou, for offering me the necessary academic background and encouragement to follow this demanding path, and to Professor Olga Maratou for her great support in one of the very many moments of crisis.

Although the Ph.D. is a solitary process, it was extremely important for me to work in the same laboratory with my very best friends Dr. Zaira Papaligoura, Dr. Geraldo Fiamenghi and Dr. Stephen Malloch. I was very lucky to have them sharing my good and bad moments. I wish themselves and their families the best of luck.

This thesis would have never been completed without the enthusiastic participation of the mothers and their babies to whom I am grateful. I would also like to thank the technicians of the Department of Psychology in Edinburgh, and the Foundation Bakalas Bros. in Greece for its financial support.

I would also like to thank my piano teacher in Greece, Dr. Efi Agrafioti, who taught me that there is no difficulty one cannot overcome.

Last but not least I would like to mention those who helped me to keep my smile and remain a human being throughout the years of this work. These are my friends in Greece, Xenia and Mara, and in Edinburgh: Maria, Emilios, Emmanuele, Rosana, Demetra, Natia, Vassilis, Aristotelis and Majeda.

This thesis is dedicated to my parents, Philippos and Magdalini, who taught me to be optimistic, realistic and patient.

ABSTRACT

Prosody, generally defined as any perceivable modulation of duration, pitch or loudness in the voice that conveys meaning, has been identified as part of the linguistic system, or compared with the sound system of Western classical music. This thesis proposes a different conception, namely that prosody is a phenomenon of human expression that precedes, and to a certain extent determines the form and function of utterances in any particular language or music system. Findings from studies of phylogenesis and ontogenesis are presented in favour of this definition. Consequently, prosody of infant vocal expressions, which are made by individuals who have not yet developed either language or musical skills, is investigated as a phenomenon in itself, with its own rules.

Recognising theoretical and methodological deficiencies in the linguistic and the Piagetian approaches to the development of infant prosodic expressions, this thesis supports the view that the origins of language are to be sought in the expressive dialogues *between* the mother and her prelinguistic child that are generated by intuitive motives for communication. Furthermore, infant vocalisations are considered as part of a system of communication constituted by all expressive modalities. Thus, the aim is to investigate the role of infant prosodic expressions in conveying emotions and communicative functions in relation to the accompanying non vocal-behaviours.

A crossectional Pilot Study involving 16 infants aged 26 to 56 weeks and their mothers was undertaken to help in the design of the Main Study. The Main Study became a case description of two first born infants and their mothers; a boy (Robin) and a girl (Julie) both aged 30 weeks at the beginning of the study. The infants were filmed in their home every fortnight for five months in a structured naturalistic setting which included the following conditions: mother-infant free-play with their own toys, mother-infant play without using objects, the infant playing alone, mother-infant play with objects provided by the researcher, a 'car task' for eliciting cooperative play, and the mother staying unresponsive. Each filming session lasted approximately thirty minutes. In order to get an insight into the infants' 'meaning potential' expressed in their vocalisations, the mothers were asked to visit the department sometime in the interval between two filming sessions and, while watching the most recent video, to report what they felt their infant was conveying- if anything- in each vocalisation.

Three types of analysis were carried out:

a) An Analysis of Prosody - An attempt was made to obtain an objective, and not linguistically based account of infant prosodic features. First measurements were obtained of the duration and the fundamental frequency curve of each vocalisation by means of a computer programme for sound analysis. The values of fundamental frequency were then logarithmically transformed into a semitone scale in order to obtain measurements more sensitive to the mother's perception.

b) A Functional Micro-Analysis of Non-Vocal Behaviours from Videos - The non vocal behaviours of mother and infant related with each vocalisation were codified without sound to examine to what extent the mothers relied for their interpretations on non-vocal behaviours accompanying vocalisations.

c) An Analysis of the Mothers' Interpretations - The infants' messages were defined as perceived by their mother.

The corpus comprised 713 vocalisations (322 for the boy and 391 for the girl) selected from a corpus of 864, and 143 minutes of video recording (64 for the boy and 79 for the girl). Correlations between the above three assessments were specified through statistical analysis.

The findings from both infants indicate that between seven and eleven months prosodic patterns are not related one to one with particular messages. Rather, prosody distinguishes between groups of messages conveying features of psychological motivation, such as 'emotional', 'interpersonal', 'referential', 'assertive' or 'receptive'. Individual messages belonging to the same message group according to the analysis of prosody, are distinguished on the basis of the accompanying non-vocal behaviours. Before nine months, 'interpersonal' vocalisations display more 'alerting' prosodic patterns than 'referential' vocalisations. After nine months prosodic patterns in Robin's vocalisations differentiate between 'assertive' and 'receptive' messages, the former being expressed by more 'alerting' prosodic patterns than the latter. This distinction reflects a better Self-Other awareness. On the other hand, Julie's vocalisations occurring in situations of 'Joint Interest' display different prosodic patterns from her vocalisations uttered in situations of 'Converging Interest'. These changes in the role infant prosody reflect developments in the infants' motivational organisation which will lead to a more efficient control of intersubjective orientation and shared attention to the environment. Moreover, it was demonstrated that new forms of prosodic expression occur in psychologically mature situations, while the psychologically novel situations are expressed by mature prosodic forms.

The above results suggest that at the threshold to language, prosody does not primarily serve identifiable linguistic functions. Rather, in spite of individual differences in form of their vocalisations, both infants use prosody in combination with other modalities as part of an expressive system, that conveys information about their motives. In this way prosody facilitates intersubjective and later cooperative communication, on which language development is built. To what extent such prelinguistic prosodic patterns are similar in form to those of the target language is a crucial issue for further investigation.

CHAPTER 1

WHAT IS PROSODY ?

The purpose of this thesis is to examine the role of the prosody of infant vocalisations in conveying messages during mother-infant communication at the second half of the first year. Before going through the literature on the development of prosodic expressions in infancy and describing this research, it is considered necessary to discuss the question "What is prosody" in order to clearly define the nature of the phenomenon to be studied.

1.1 A History of the Concept

The term 'prosody' is directly derived from the ancient Greek word 'prosodia' (προσωδία), which was originally used in ancient Greek literature to mean a song accompanying instrumental music (Liddell and Scott, 1882).

Plato introduces a pioneering approach to the definition of prosody. In the *Republic* (Book III, X) the philosopher states that a song is composed of three parts, the words, the rhythm and the tune. The tune is the 'harmony' that is, a pitch system of high and low notes. Harmony is used technically in the scales and modes of Greek music. Plato, representing the majority beliefs of the time, claims that the musical modes are adapted to the expression of particular feelings and spring from the prosody (προσωδία) of human speech i.e., the variation of pitch of the speaking voice when it is expressing similar feelings.

In Aristotle (*SE* 166b I, 177 b 3) 'prosodia' referred to modulations in pitch of the speaking voice and to the pronunciation of a syllable on a certain pitch. Later scholars include in the meaning of prosody other pronunciation characteristics as well such as syllable quantity and vowel breathiness (Liddell and Scott, 1882).

The scholar Herodianus in his famous treatise "*On Prosody*" (=Περί Προσωδίας), written in the second half of the 2nd century AC, gives a thorough and precise definition of the word: "Prosody is any certain kind of modulation of every writable logical voice". An explanation of the above definition is found in the "*Comments on Dionisius Grammar*" (=Σχόλια εἰς τὴν Διονυσίου Γραμματικήν) (Bekker, 1816 pp. 676-7): "Prosody is any voice modulation.....of a certain quality; this modulation is either tense or lax or medium. There are two kinds of sounds: the

articulated sounds which can be written.....and the inarticulated ones that cannot be written,.....but because even inarticulated sounds can be modulated as for example the ones produced by knocking on iron or wood, which cannot basically be called voices, for this the characterisation 'writable' was added in the definition. Moreover, there are writable voices that do not convey any meaning, for this reason the characterisation 'logical' was added in the definition describing a voice that conveys a meaning. One must know that the term 'prosody' does not only include the specified pitch heights.....but also the vowel length and breathiness, since the length can also be modulated i.e. it can be shorter or longer".

Furthermore, prosody refers to the succession of tonally accented longer and shorter syllables; this is a fundamental element of pronunciation in ancient Greek and was determining the rhythm in ancient Greek poetry. Different succession patterns were used by the ancient Greek drama poets for the narrative and the dialogic parts of a play (Lupourlis, 1978). Later, scholars of the English poetry such as Ambercombie (1923) used the term 'prosody' to refer to the metrical organisation of the English poetry, which was based on the succession of dynamically stressed and unstressed syllables.

Burnet (Lord Monboddo) (1773-92) asserted that there are syllabic accents in English, but there is no change of tone in them; the voice is only raised more, so as to be louder for one syllable than for another. "The music of our language in this respect is nothing better than the music of a drum in which we perceive no difference except that of louder and softer". Contrasting Burnet, Steele in his treatise *Prosodia Rationalis* (1775) demonstrated the existence of tonal variation in English and provided the first systematic transcriptional method for notating what he considered to be prosodic features namely, 'length', 'stress' and 'pitch' features. He noted that, although not all pitch variations are significant, gradations between tones are too rapid and small to be described by the diatonic and the chromatic genii of music. The scholar also discusses both the grammatical and the attitudinal function of pitch.

In the early definitions of prosody described above is obvious that the term is related to the two most characteristic cultural forms of human communication, language and music.

1.2 Linguistic Definitions of 'Prosody'

A number of linguists have subsumed different features under the term 'prosody' and identified the linguistic status of prosodic phenomena in various ways. These can be classified as of three trends: a) Prosodic Elements are Determined by other Parts of the Linguistic Structure, b) Prosodic Elements are Part of the Linguistic Structure of Equal Importance with the other Linguistic Systems, and c) Prosody is a Marginal Element of the Linguistic System.

a) Prosodic Elements are Determined by other Parts of the Linguistic Structure

De Angulo (1929) suggests that the term 'prosody' should include all those changes in 'duration' (or 'length', 'quantity'), 'pitch' (or 'tone') and 'loudness' (or 'volume', 'stress') that are grouped together as 'accentual differences'. These changes are essentially grammatical processes driven by grammatical rules to express grammatical concepts. Thus, according to this author, prosody is completely subordinated to grammar.

Bloch and Trager (1942) claimed that over and above isolated vowels, or consonants or successions of these, are superimposed particular variations in the length, voice pitch and loudness. These variations constitute the prosodic features of 'quantity' (length), 'tone' (pitch) and 'stress' (loudness); the last two are usually grouped together forming the 'accent'. When realised over syllables and words, the prosodic features are referred to as 'accents', and when realised over phrases and sentences they are described as 'intonation'. However, in another part of their work the authors admit that prosodic features constitute as much an integral part of the utterance as segmental sounds. Bloch and Trager (1942) were the first to integrate the notion of prosodic features into a general phonological theory and to suggest a methodology for their analysis, which involves segmental description in the first place and then prosodic or suprasegmental description.

In the same vein, Haugen (1949) used the term 'prosodeme' to describe modifications in 'duration', 'pitch', 'stress' and 'juncture'. These modifications depend strongly on the timing of the syllable. Thus, according to Haugen, a valid analysis of prosody cannot be done without some implicit or explicit definition of the syllable. For Meyer-Eppler (1957) the term 'prosodic' refers to features belonging to a sentence as a whole and expressed by 'pitch' and 'stress' patterns.

b) Prosodic Elements are Part of the Linguistic Structure of Equal Importance with other parts of the Linguistic System (i.e. segments, vocabulary, grammar).

Crystal asserts that prosody is a phenomenon that can be described only in elements which manifest phonological contrastivity and, thus, should be considered as an aspect of the linguistic structure (Crystal, 1969, 1973a). The subject-matter of prosodic analysis is the system of linguistic contrasts in the non-segmental 'residue of utterance', after one has subtracted from speech the aspects of segmental phonology, vocabulary and grammar. In other words, the prosodic systems are sets of mutually defining phonological features that have an essentially variable relationship to the words selected, as opposed to such features as the segmental phonemes and the lexical meaning, which have a direct and identifying relationship to such words. The author also distinguishes the 'prosodic' from the 'paralinguistic' effects; although both are non-segmental aspects of speech the former are primarily the result of activity of the vocal cords', while paralinguistic effects are primarily the result of other physiological mechanisms located in the pharyngeal, oral or nasal cavities. The 'paralinguistic' effects include the 'voice quality' (e.g. whisper, creak) and 'voice qualifications' (e.g. laugh, cry) (Crystal, 1969).

Crystal (1969, 1973a) sees prosody as having hierarchical organisation and he considered as primary prosodic parameters those non-segmental characteristics of speech which refer to variations in such psychological attributes of sound as 'pitch', 'loudness', 'duration' and 'pause'. 'Pitch direction' and 'pitch range' constitute two of the fundamental contrastive systems of prosody. A higher level of prosodic organisation includes 'tempo' and 'rhythmicality'; these prosodic systems comprise independently varying vocal effects that are based on combinations of the primary parameters in specific ways. More specifically, 'tempo' refers to the rate of speech, and is depended on the timing of syllables, the occurrence of pauses, and the length and syntactic complexity of larger grammatical units. Tempo has probably the most highly discrete grammatical function of all prosodic parameters other than pitch. 'Rhythmicality' is the prosodic system which accounts for those linguistic contrasts attributable to our perception of regularly occurring peaks of prominence in utterance. The most general level of organisation that can be imposed upon prosodic features is described by the 'tone unit'. A 'tone unit' (or 'primary contour' or sense group') is a configuration of features of 'pitch direction', 'pitch range', 'rhythm' and 'pause'. A 'tone unit' is equivalent in status to the notion of 'sentence' in grammatical analysis. The assignment of tone unit boundaries seems motivated by syntactic reasons.

Further, prosodic feature carrying important linguistic contrastivity is the placement of maximum prominence on a given syllable, or occasionally on more than one syllables, in the tone unit (called 'prominent' or 'tonic' syllable). This prominence primarily results from a pitch movement of certain direction and range , but 'loudness', 'duration' and 'pause' may also be controlled to heighten the contrast between what precedes and follows (Crystal, 1979).

Moreover, under the term 'prosody' the author subsumes 'intonation', which refers to a system of linguistic contrasts comprised mainly of 'pitch height' and 'pitch range' and secondarily by 'rhythmicality' and 'tempo' (Crystal and Quirk, 1964; Crystal, 1969,1973a, 1979). The point at which pitch contrast becomes completely subordinated to vocal or non-vocal effects of a different nature is the point at which intonation gives way to other communicational systems (Crystal, 1969).

In this framework Crystal (1969, 1975; Crystal and Quirk, 1964) claims that prosody is in fact an essential part of the linguistic system since it fulfils the following criteria: 1) Its features of articulation are conventional that is, determined by cultural tradition rather than by human physiology alone; 2) these features are language specific, i.e. differing in pattern and realisation from one language to another; 3) they are part of a code that is, different patterns convey contrastive meanings attached to them conventionally and arbitrarily; 4) they show internal pattering and contrastivity and, thus, they are capable of phonological analysis in structural and systemic terms, although in fact scales of contrastivity are manifested; 5) they are identified in terms of their phonetic shape, which constitutes discrete units, allowing for the fact that different features may have different degrees of discreteness; 6) prosody manifests integrability with other aspects of linguistic structure, particularly grammar; 7) it is hierarchical; and 8) it works in sequence or simultaneously.

Other authors (Barry, 1981; Couper-Kuhlen, 1986; Cruttenden,1986) also allow for three fundamental prosodic features, which are most consistently used for linguistic purposes, either singly or jointly; namely, 'length', 'pitch' and 'loudness'. Like Crystal they also include in the definition of prosody elements produced by the combination of the fundamental features and characterised as linguistic in nature, such as word and sentence 'accent' (or 'stress'), rhythm, intonation, tempo and pause. Cruttenden (1986) calls pauses 'paralinguistic' effects, since, unlike the other prosodic features, they are interruptive of speech rather than co-occurrent. In opposition to Crystal, the same author considers 'voice quality' as a prosodic feature. Couper-Kuhlen (1986) also identified prosody negatively, as those auditory components of an utterance that are not

segments, not idiosyncratic (e.g. individual's voice quality), not physiologically determined and not 'paralinguistic'; that is, effects that are only sporadically present in the speech, and they are not systematic conventional or intentionally used in communication. Prosody comprises continuously present effects in speech; it is not normal to make utterances without such effects (Couper-Kuhlen, 1986).

c) Prosody is a Marginal Element of the Linguistic System.

Welmers (1959) uses the term 'para-linguistic' to cover the linguistic as well as the non linguistic functions of 'duration', 'pitch' and loudness, whilst Brown (1977) under the same term includes certain features of pitch such as 'pitch height' and 'pitch range'. Martinet (1964) states this viewpoint more explicitly. He notes that 'prosody' refers to all the features of speech that do not fall within the phonematic framework, and he characterises prosody as a linguistically marginal phenomenon, since a feature is considered as properly linguistic only in so far it is analysable into phonemes. Although he attributes to prosodic features some linguistic value, this is different from the linguistic value of phonemes, in that prosodic features signal meaningful contrasts not by their mere presence or absence, since they are always present, but by their modifications. The author concludes that linguistic research could show regular prosodic form-function correspondences.

Some linguists attempt to define the functions of prosody in communication. In their approach Carrel and Tiffany (1960) propose that the prosodic elements of speech are 'duration', 'pitch', 'strength' and 'quality' and that their function is to carry meanings over and above the bare intellectual content of the words. However, as Crystal (1969) noted an absolute distinction cannot be drawn between 'intellectual' and non-intellectual or emotional content. Rather, a prosodic pattern conveys different kinds of messages simultaneously, some of which become dominant in the situation where the particular pattern is uttered. Garnica (1977) uses the terms 'prosodic' where the features such as 'pitch', 'loudness' and 'stress' serve the marking of lexical and grammatical differences, and the term 'paralinguistic' where the same features mark attitude or emotion.

In studies of the melody in infant vocalisations the terms 'prosody' and 'intonation' have been used interchangeably (Crystal, 1979; Vihman, 1985, 1986; Marwick, 1986; Marcos, 1987; D'Odorico and Franco, 1991). This is attributable to the fact that these authors considered only the possible linguistic aspects of infant vocalisations and did not treat them as an expressive system in its own right.

Laver (1994) introduces a new approach to the concept of prosody by claiming that although the prosodic elements of duration, pitch and loudness contribute to the formation of the linguistic structure, they are themselves not linguistic in nature, since they do not necessarily require the phonological basis of syllables, words or phrases for their realisation. So these elements should not be subsumed under the same term with 'intonation' and 'stress', which represent the linguistic use of pitch and a phonological property of the syllable, respectively. Moreover, although prosody is considered to be one of the three patterns of segmental organisation, together with speech 'tempo' and 'rhythmicality', it is still to be distinguished from these two, which depend on phonological features such as 'syllabic timing', 'stress', 'syllable weight' and 'articulation rate' (Laver, 1994).

The prosodic organisation of any individual utterance is formed by variations in time of pitch and loudness, and is generated by successive settings of the vocal tract. These settings are constituted by any tendency for the vocal apparatus to maintain a given configuration or featural state over two or more segments in close proximity in the stream of speech (Laver, 1994).

Laver (1994) suggests that the analysis of prosody of a particular utterance should involve a description of the 'melodic pattern'; that is, the dynamic pitch movement or 'pitch contour', which refers to the 'shape', 'direction' and 'variability' of the trajectory manifested by any perceivable change in pitch height realised over syllables, words or phrases, and the 'pitch span', which is the local range within which the speaker organises relative values of pitch in order to form the melodic contour of an utterance. Similar features should also be described for the parameter of loudness.

In a discussion of prosodic features, one needs to distinguish these from the so-called inherent features of sounds. A segment is defined by the mutual co-presence of particular features, including duration, pitch and loudness. On the other hand, the prosodic features are independent of the definition of segments (Jackobson and Halle, 1956; Lehiste, 1970; Laver, 1994).

At this point it is necessary to recall Haas's (1970) warning that the term 'prosody' should not be used as synonymous to the term 'suprasegmentals', as it has been some authors (Pike, 1945; Bendor-Samuel, 1960; Luchsinger and Arnold, 1965; Fromkin and Rodman, 1970). 'Suprasegmentals' are defined as those features that are 'overlaid' on a segment, and that can also potentially be realised beyond the domain of a segment. They are not merely modifications in 'duration', 'pitch' and 'length', but involve also the

phenomena of vowel length, sound reduction, elision, coarticulation, assimilation, dissimilation and vowel harmony, which extend over more than one segment (Lehiste, 1970; Couper-Kuhlen, 1986). Crystal (1969) also finds the term 'suprasegmental' unsatisfactory as a substitute for the term 'prosody', because the prefix 'supra-' implies a priority of segmental over non-segmental linguistic features.

1.3 The relation between Prosody and Music

A number of authors have emphasised the relation between prosodic phenomena and music. Spencer (1858) and Darwin (1877) characterise as 'musical' any sound capable of modulation in 'pitch', 'loudness', and 'quality' or 'resonance and 'timbre', uttered by any animal to express emotions. Luchsinger and Arnold (1965) proposed the distinction between 'verbal' i.e., morphemic and phonemic elements, and 'additional' elements of communication. The latter constitute the 'musical basis' of language, which includes the 'prosodics', namely 'duration', 'pitch', and 'loudness'. Similarly, Fonagy and Magdics (1963), in referring to the 'musical' or 'prosodic' elements of speech specify 'intonation', 'stress', 'speech-tempo' and 'pause'. Bolinger (1964) observes that when we speak, the fundamental frequency, which is the physical correlate of the perceivable attribute of pitch, is used in a manner very similar to that of ordinary music. The author calls the use of fundamental frequency in speech 'intonation'. Intonation resembles music not only in its physical basis but also in that both convey emotions (Bolinger, 1964). In their study of infant vocalisations the Papouseks (1981) use the term 'musical elements' instead of the term 'prosodic elements' to describe the voice 'pitch', 'loudness', 'tempo', and 'rhythm'. The authors justify this preference by pointing out that in studies of preverbal vocalisations it is difficult to make linguistic attributions. Furthermore, they want to pay attention to both linguistic and musicological aspects of the infant vocal behaviour.

In order to properly search for any relation between 'prosody' and 'music', one needs in the first place to give a precise definition of both. The above authors who claim similarities between these phenomena, although they do not explicitly state it, seem to refer by the term 'music' to the whole of compositions of one particular musical system, namely, Western music as an art form. Although these authors observe a relation between prosody and a musical system, they do not specify this relation. Moreover, there is no agreement among the scholars about what is subsumed under the term 'music'. Many have borrowed terminology from music in their search for more precise and economical ways of describing prosodic features. In turn, certain

composers such as Berio and Stockhausen have used combinations of linguistic and musical notation in their pieces for voice.

Further speculation on the relations between music and speech in phylogenesis goes beyond the scopes of this thesis, but reference is made to musical descriptions of infant vocalisations in Chapter 3.

1.4 The Functions of Prosody

The several features described by different authors as prosodic serve a variety of functions. In tone languages (e.g. Mandarin and Cantonese Chinese, or Zulu) the meaning of words that are identical in phonemic structure is differentiated by changing the pitch level at which they are uttered. This is the lexical function of prosody. In a grammatical function, changes in pitch level distinguish between tenses in some languages (e.g. Bini and Twi in West Africa). Also the prosodic element of pitch may be used to identify different units in the chain of language such as sentences and phrases (Crystal, 1969, 1979; Laver, 1994).

In the semantic function, prosody indicates the status of an utterance within the framework of discourse, relating utterances to one another, reflecting presuppositions about the discourse and emphasising, where necessary, certain lexical items (Bolinger, 1989; Crystal, 1979).

The psychological function of prosody is evident from several experiments which have shown that performance in short-term memory, recall, perception and other variables is affected by the prosodic character of an utterance. For example, it has been reported that words containing the most stressed syllables in an utterance are more readily perceived, attended to, and recalled (Blasdell and Jensen, 1970).

Prosodic features also signal information about the sociolinguistic characteristics of the speaker, such as gender, class or professional status. This is the social function of prosody (Crystal, 1969, 1979). Under this category Crystal (1969) also includes what he calls the 'conversational device' function of prosody, which facilitates the turn-taking in conversation. As it will be explained in the second chapter this intersubjective function is already paramount in 'protoconversations' between mothers and their two month old infants (Trevarthen, 1982). Cruttenden (1986) calls these co-occurrent or interruptive effects that have no conventional meaning, and that are conditioned by factors over which the speaker has no immediate control, 'extralinguistic'.

Pivotal for communication is the attitudinal function of prosody (Crystal, 1979). The phonetic character of pitch behaviour, especially as a consequence of its gradient nature, can be used effectively to convey subtle features of the speaker's emotions or attitudes concerning the subject matter or context of an utterance. Laver (1994) calls this function 'paralinguistic', while Cruttenden (1986) considers that as 'linguistic', and 'prosodic'. It is unclear to what extent such use of prosodic features to express emotions can be specific to a language, or expressive of universal characteristics of emotion (Bolinger, 1964).

At one point Crystal (1979) claims that the primary role of prosody is the expression of the 'illocutionary force' of a 'Speech Act', such as 'persuading' or 'commanding', though whether one might refer to this as primarily a social, attitudinal, semantic or grammatical role is very much an open question. Moreover, Crystal (1969) makes the crucial remark that there is no one-to-one correspondence between prosodic forms and prosodic functions, and Marwick (1986) further argues that in any conversational utterance at least the 'semantic', the 'attitudinal' and 'social' functions of prosody must be expressed simultaneously. In her study of mother and infant intonation in the second year of life this author defines the **interactive** function of prosody as the one which incorporates information from all the other functions.

1.5 Theoretical Framework for studying Prosody in Infant Vocalisations

Taking into account the points made in the preceding literature review, I would define prosody operationally as any perceivable modulation of duration, pitch or loudness which conveys meaning, and I would put forward the argument that, in this sense, prosody is a phenomenon preceding, and to a certain extent determining, any particular language or musical system, in terms of both form and function. Regarding the parameter of form, it is the case that modulations in 'duration', 'pitch' and 'loudness' can be manifested in any utterance produced by vocal cord activity and showing periodicity. As to the function of vocal behaviour, whereas verbal communication of ideas is considered exclusive to human society and exists in many specific variations, 'musical vocalisations' are common to a large part of the animal world and they serve many purposes in social groups (Papousek and Papousek, 1981). The arguments in favour of this position come from data on the phylogenesis, the role of maternal speech to infants ('motherese'), as well as from the ontogenesis (both the perception and the production) of vocal behaviour.

a) Arguments from the Phylogenesis

In his monumental treatise *The Expression of Emotion in Man and Animals* (1877) Darwin observed that in most of the social animals, including humans, variations in loudness, quality and pitch communicate the need for help, pleasure or anger, however these may be aroused. Seeking the origins of this ability, Darwin argues that the progenitors of man probably uttered 'musical tones' to transmit different emotions before they were able to utter articulated speech. He believed that the habit of uttering 'musical sounds' was first developed as a means of courtship in the early progenitors of man and, thus, became associated with the strongest emotions namely, ardent love, rivalry and triumph. Consequently, through the principle of association, when the voice is used under any strong emotion, it tends to assume a 'musical character' (Darwin, 1877).

In the same vein, in his study of what he called 'living primitive languages', Jespersen (1922) suggests a common evolutionary origin of both singing and verbal language in a 'protolanguage' characterised by high proportion of 'musical elements' that is, 'melodic contours' and 'pitch accents'. In this sense 'human protolanguage' was regarded as closer to singing than to speech (Jespersen, 1922) Following Darwin's (1877) speculations, Jespersen (1922) also claimed that the musical elements of speech serve mainly the expression of emotions rather than the communication of ideas, and considered this fact to be a crucial factor in the development of language. The author described an atrophy of prosodic elements increasing with the evolution of human languages to highly abstract forms.

b) Arguments from 'Motherese'

The most robust support for the idea that human vocal prosody is a phenomenon preceding language in terms of form and function, rather than one defined through language comes from studies on 'motherese' that is, the speech with which mothers address themselves to infants. It has been stated that "mothers' speech to infants has the part of prosody particularly strong" (Papousek and Papousek, 1981 p. 169). 'Motherese' is characterised by high average pitch level, smoothly gliding contours, expanded pitch excursions, greater pitch variability, short utterances and longer pauses, and simplified syntactic structure (Sachs et al, 1976; Papousek and Papousek, 1981; Stern et al, 1982; Papousek, Papousek and Bornstein, 1985; Fernald et al, 1989).

Regarding the function of 'motherese', Ryan (1978) proposed that the communicative function at six months of a rising contour in mother's speech to her infant conveys request for visual attention, whilst at one year the same pattern expresses invitation to join in or contribute in some way to the interaction.

Stern, Spieker and McKain (1982) analysed the utterances of six mothers to their 2, 4, and 6 month old infants to determine the type of grammatical sentence, the pitch contour as well as the interactional and motivational context, in which these utterances occurred. This study demonstrated that mothers used specific pitch contours in specific interactional contexts. In particular, in context 1 (infant gazing away from mother and showing neutral affect) the mother's main intention appeared to be to get the infants attention, and rising contour was used to convey this function. In context 3 (infant displaying positive affect and gazing at the mother) the mother's main purpose seemed to be the maintenance of the infant's positive affect; this purpose was served by the bell contour (rise-fall). Finally, in context 2 (infant gazing at the mother with neutral affect) the mother's intention was generally to get more intense engagement, this was expressed by various contours. Furthermore, Stern and his colleagues (1982) note that mothers used certain pitch contours with different sentence types, analysed purely on the basis of grammatical form. Specifically, 'Yes-No' interrogative forms were found to have rising contours whilst 'wh' interrogative forms and imperative forms in general displayed falling contours. The declarative forms had bell-shaped contours. These contour shapes appear in the English intonation system with these grammatical forms.

One might argue that the relationship demonstrated between contour shape and communicative context is secondary to the fact that mothers used certain sentence types in certain contexts, since it has also been shown that specific sentence types are associated with particular contours shapes. This argument, however, is moot as to whether the contour-in-context finding resulted from greater use of sentence types with a particular contour in a given context, or from the changing of a sentence's contour depending on the context in which it is spoken. However, 2 to 6 month old infants clearly cannot discriminate sentence types linguistically; on the contrary they are able to grasp intersubjective meanings through prosodic contours months before language becomes linguistically meaningful. So, from the infant's point of view it does not matter what linguistic or other means the mother uses to assure that certain contours will be produced in certain communicative contexts (Stern et al., 1982).

Papousek and her colleagues (1991) analysed form-function relations between melodic contours in maternal speech and interactional caregiving contexts specified in

spontaneous interactions between Chinese and American mothers and their two month old infants. It was demonstrated that the mother's use of contour types was influenced by the communicative context, and only minimally by language content. Specifically, rising contours were systematically used by mothers in contexts where the mothers encouraged the infant's active participation in the dialogue, whilst falling or bell shaped contours mainly appeared in contexts where the mothers responded to infant behaviours. Thus, mothers seem to mark their turn-opening and turn-closing interventions, establishing the basic pattern of dialogue, i.e. the turn-taking with rising versus falling or bell contours. Other characteristics of the pitch contours, such as pitch height, pitch range, slope and duration, were also significantly linked to particular caregiving contexts.

There is increasing evidence that 'motherese' is an intuitive and universal behaviour by which mothers enter into intersubjective relation with their infants. 'Motherese' is characterised as 'intuitive' because its features require no thought or conscious preparation (Trevarthen, 1990; Trehub, 1993). The patterns of prosodic modification in adult speech to preverbal infants show remarkable consistency in French, Italian, German, Japanese, and British and American English (Ferguson, 1964; Fernald and Simons, 1984; Grieser and Kuhl, 1988; Fernald et al., 1989) Also, it has been reported that American and Chinese mothers use the same contour types to convey the same kinds of meaning in relation to the majority of contexts (Papousek et al, 1991). What is striking about these findings is how Chinese mothers who produce more complex fundamental frequency patterns with higher rate of fundamental frequency change in adult directed speech (Eady, 1982) gave to their infants the same simplified melodic prototypes as American mothers. It was shown that Chinese mothers avoided lexical utterances in favour of interjections, exclamations, matching and modelling sounds just as do American mothers, or German mothers and fathers (Papousek et al, 1987). In their lexical utterances, Chinese mothers seemed to employ a number of intuitive strategies which allowed them to outwit the tone rules of their language. In some cases, they even violated tone rules and drastically changed lexical tones in favour of prototypical caregiving melodies (Papousek and Papousek, 1987).

The studies on the meaning of parental melodies in spontaneous communication with their prelinguistic infant in many different cultures and linguistic communities described above, provide evidence that in human ontogeny the prosodic elements of pitch and duration are modulated systematically by the mother in a non-language-specific way in order to facilitate communication with an infant. Maternal melodic contours act as a stable information carrying unit expressing the mother's motives,

emotions and intentions which match with the infant's motives, emotions and intentions in a communicative interchange (Stern et al, 1982). Further, maternal prosody is modified as the infant grows older in order to transmit novel messages which will meet the new requirements for communication generated by the new motives active in each developmental stage (Trevarthen and Marwick, 1984). Of course there are cultural differences because of the different cultural conventions in the non verbal display of attitudes and emotions; however, these differences concern the extent and not the quality or intersubjective message of melodic adjustments (Papousek et al, 1991).

c) Arguments from the Ontogenesis

There is increasing evidence that infants are inherently adapted to attend to, differentially recognise and respond to the messages conveyed in maternal prosodic patterns as functional units. Hutt and colleagues (1968) showed that infants are sensitive to the frequency range in which mother's utterances are likely to be spoken. Moreover, it has been found that three weeks old infants were capable of discriminating their mother's normal voice, but not if she was reading the text from right to left or up and down and therefore with abnormal prosody (Mehler et al., 1978; Mehler et al, 1979). The authors concluded that this ability is primarily based upon the prosodic features in the mother's speech, which was probably learned before birth. Work by De Casper and his colleagues has postulated that prosodic aspects of maternal speech such as rhythm and pitch modulations are audible to the foetus in utero, and that this may contribute to, or even determine, the very early preference observed in newborns for their mother's voice (De Casper and Fifer, 1980; De Casper and Spence, 1986). This preference has also been demonstrated for four month old infants by Chiang (1992). The four-month-olds preference for infant directed speech persists even when the lexical content, but not the prosody is removed by filtering (Fernald, 1985).

Further, Crystal (1979) suggested that two months old infants are aware of prosodic contrasts in adult utterances directed to them, and more specifically Morse (1972) had shown that at this age infants are able to discriminate 'rising' and 'falling' contour shapes. Two- to four-month-old infants showed an ability to discriminate categorical but not arbitrary prosodic shifts independently of language differences (Best et al, 1982). Menyuk and her colleagues (1986) provided evidence that after six months infants discriminate and categorise different prosodic patterns in terms of communicative intent. It would be interesting to compare findings on infant perception of segmental and prosodic features. However, as Best (1992) has noticed, such

comparison is problematic, because the former have focused on discrimination of minimal contrasts, whilst the latter have largely tested preference for global patterns.

Wolff (1969) reported that young infants respond differentially to differences in voice pitch. Five- to nine-month-olds were shown to respond to familiar prosodic contours with situation-specific behaviours (Kaczmarec cited in Weir, 1966). At six months an infant is able to discriminate between an angry and a friendly tone of voice and react accordingly (Lewis, 1936; White, 1982). Lieberman (1967) reported this behaviour from the third month. Studies of the kinesics of early mother-infant interactions have demonstrated an ability in the newborn to synchronise its motor activity with the rhythm of adult speech (Condon and Sander, 1974). Infants have been shown to respond to motherese with temporally synchronised coordinated body, facial and vocal expressions (Brazelton et al, 1979).

In a series of five auditory preference experiments carried out by Fernald (1993), 120 American, five-month-old infants were presented with Approval and Prohibition vocalisations in infant directed and adult directed English, and in infant directed speech in nonsense English and three unfamiliar languages, German, Italian, and Japanese. Dependent measures were looking-time to the side of stimulous presentation, and positive and negative facial affect. The finding that five-month old, American infants show more positive affect to Approvals and more negative affect to Prohibitions, spoken in two unfamiliar languages i.e., German and Italian as well as in nonsense and natural English, provides evidence that young infants respond differentially to positive and negative messages conveyed only by prosodic features without facial affect.

Studies on the prosodic characteristics of prelinguistic vocalisations have also revealed that these include features which also appear in conventional musical systems or poetry. Analysing the development of rhythmic vocalisations, Fridman (1980) observed that a variety of 'protorhythms' is already obvious in the newborns' expressions. The author has considered those 'protorhythms' as 'natural rhythmic schemes' underlying the rhythms of both adult speech in various languages and Western classical music.

The Papouseks (1981) reported that in their daughter's vocalisations all common intervals of pitch appeared in the period from birth to sixteen months. However, in all age periods the majority of them was unisons and seconds (67%) followed by fourths, minor thirds, perfect fifths and octaves. These intervals occur frequently in traditional nursery songs, and according to musicologists they belong to the most frequently sung

intervals all over the world (Nettl, 1956; Bernstein, 1976). Furthermore, the Papouseks note that from the beginning of life the infant's vocal rhythm seems to be closer to folk music or jazz improvisation than to the Western classical music.

Lynch and colleagues (1994) claimed that 'phrasing'; that is, the temporal organisation of behaviour based on the duration of its segments which is a fundamental prosodic feature, is an innate and universal characteristic of human communication. The authors explored the developmental roots of phrasing in non-vegetative, prelinguistic vocalisations in a longitudinal study of eight typically developing infants and eight infants with Down syndrome, between two and twelve months. Adult judges identified hierarchical arrangements of syllables embedded within utterances and utterances embedded within phrases. Prelinguistic phrases were characterised by systematic lengthening of phrase-final syllables, temporal patterning, and stable durations across development; but more important, the temporal structure observed was similar to that of some cross-culturally optimal rhythmic units from other domains. The mean syllable duration in infant vocalisations corresponds to that of adult syllable duration (mean value 350 milliseconds), the utterance duration to the metrical foot in poetry (mean value 500 milliseconds) and the phrase duration to the poetic line or the musical phrase (mean value 4 seconds). Analysis of vocalisations of infants with Down syndrome indicated similar internal structure of prelinguistic phrases to those of typically developing infants, but with slightly longer durations, and under larger range of variation.

The above discussion presents data from the phylogenesis and the ontogenesis supporting the view that, in terms of form and function, prosody comes prior to any conventional linguistic or musical system and may constitute a generative matrix for both these cultural means of expression, rather than being determined by those. Thus a study of infant vocal expressions can inform about foundations for the whole range of human communication in all its applications and cultural adaptations.

CHAPTER 2

PREVIOUS STUDIES ON INFANT PROSODY.

2.1 Early Research

As early as 1876 Taine reported that his child was able to express all the shades of emotion, wonder, joy, wilfulness and sadness by differences in the tone of her voice, in a richer manner than an adult. Moreover, compared even to the most gifted animals in this respect, such as parrots and singing birds, this infant's ways of expressing intonations were far more delicate and abundant, although the gamut of sounds bearing them was less extended.

Charles Darwin was the pioneer in the study of prelinguistic vocal behaviour. His naturalistic observations of his own infant and detailed descriptions provided the basis for further research in the field. Darwin (1877) noted that at the beginning of life infants express their needs by instinctive cries. After some time these cries are modified in part unconsciously and in part voluntarily as a means of communication. When one and a half months old, his son first made little noises without any meaning, to please himself, and these soon became varied. When a little over a year old, the same infant raised considerably the 'musical pitch' of his voice when pronouncing the word 'mum' while asking for food. Moreover, he produced an exclamatory sound when recognising a person or his own image in the mirror. Darwin notes that the same melodic patterns are used in similar circumstances by adults, but he does not attribute this phenomenon to any learning process. Rather, he claims that the use of these 'intonations' seemed to have arisen instinctively. As Crystal (1973a) notes 'intonation' in Darwin's papers means the general melodic impression obtained from listening to infant vocalisations. Under the age of two years, the baby was expressing his triumph of assent by a slight modulation strongly emphatic, and obstinate determination by a 'peculiar whine' (Darwin, 1872). Compared to Taine's, these descriptions are of a more functional nature. Summarising his viewpoint on infant vocal development Darwin (1877) asserts that the infant's needs at the beginning are expressed in cries, later by gestures and in a marked manner by intonations, and lastly by words of a general nature invented by himself, then of a more precise nature imitated from those whom he hears. These observations support for ontogenesis what Darwin claims to be also true in phylogenesis concerning vocal expression i.e., this is realised at first as "notes in a true musical scale" (Darwin, 1877 p. 293) and then elaborated in an articulated form.

Jespersen's work is also considered to be fundamental in the study of infant vocalisations. In his classical essay "Language: Its Nature, Development and Origin" (1922) he distinguishes three periods in language development. During the first period characterised as the 'screaming time' infants express themselves mainly by loud high pitched noises. It is noteworthy that the author does not use the term 'cry'-which becomes very popular later in the literature. Ostwald (1963) pinpoints a qualitative difference between 'cry', which has 'tonal quality' and 'scream', which is characterised by 'rasp'. The second period described by Jespersen (1922) is the 'babbling time', when infants utter sound series without any referential meaning, for the sake of delightful exercise. The author suggests that these sounds are phonemically different from the adult ones, and even if they resemble those, the mechanism of production is different. It is characteristic of the nature of these sounds that at least some are difficult for the child to produce when real words appear, and may be totally unknown to the target language. However, in this repertoire coexist the germs for the development of both singing and verbal language, although the child's repertoire at this age is closer to singing than to speech. Jespersen emphasises that the prosodic features of these vocalisations express emotions, and he considers this compound to be a crucial factor in language development. Later in this period, some of the prosodic elements in infant early vocalisations now develop, by an innate mechanism, into linguistic prosodic patterns conveying meanings; this is the first aspect of the adult language to be acquired. Jespersen, however, neither describes nor interprets adequately the transitional phenomena. The third period of language development is named 'talking time'. At the beginning children produce 'little language'; then, at some point they know what sound they wish to utter and are able to produce it exactly. Thus, they eventually acquire the language of the community. The author emphasises that the order of language acquisition described above mirrors the natural of the aspects of language hierarchy.

In 1931 M. Fitchen wrote an article entitled "*Speech and Music Development of One Year Old Child*". This is the first time that explicitly stated aim to examine the development of speech and musical ability in combination. Fitchen observed that her subject, M.H., when almost one year old could speak comparatively clearly with all the vocal sounds present, although the final consonants were usually omitted. 'Musically', the sounds of M.H. did not at first vary much either in 'pitch' or in 'tone'. For the most part they ranged over the eight notes of the octave from middle C up. Due to the nature of her vocabulary, M.H.'s speech is almost entirely staccato, and therefore little comparison can be made between her speech and rhythm. The individual words are

tonal and pure. The infant showed a favourable and appreciative attitude towards music and rhythm of different kinds. Her spontaneous singing was quite melodious, but not usually very rhythmic.

Although these early accounts of prelanguage vocal development are scattered and based on impressionistic data (Crystal, 1973a), they are striking in suggesting that infants are endowed with a prosodic expressive means which vocally contributes to the expression of emotions or communicative functions, long before they acquire the adult phonemic system. These prosodic features are not linguistic; rather they appear constitute a generative matrix for the development of the linguistic prosodic patterns.

2.2 Recent Studies on Infant Prosody

Scholars of prosodic features in infant vocalisations have focused on different aspects of the issue, generating different research traditions.

(a) Studies on Age-Related Modifications of Acoustic Features

These studies aimed to detect developmental changes in certain acoustic features of the infants vocalisations presumed to result from maturation of the vocal tract.

In a longitudinal study of two infants, a boy and a girl, Sheppard and Lane (1968) investigated changes in prosodic features of cry and non cry vocalisations during the first five months of life. Both infants show a decrease from their initial fundamental frequency values during the first twenty one days, then an increase to a value that exceeds the initial level, and finally stability for the remaining duration of the study. The mean fundamental frequency values were higher for the boy (433Hz) than for the girl (402Hz). Also, the infants' utterances did not change in fundamental frequency or amplitude variability with age. However, within-utterance variability in amplitude was greater than within utterance variability in fundamental frequency. Regarding the parameter of duration, the mean value for both infants was 550 milliseconds; it was found that there were as many short utterances as long ones. However, the average duration of utterances within a sample became more uniform with increasing age.

Laufer and Horii (1977) examined fundamental frequency characteristics of non distressed vocalisations produced by four infants during the first twenty four weeks of life on samples obtained from home and laboratory recordings. The results on fundamental frequency level and duration are somewhat different from those of

Sheppard and Lane (1968). The mean fundamental frequency for pooled data remained fairly stable during the period of study around 335Hz. On the other hand, developmental changes were noted in the features of duration and within-utterance range and variability. In particular, in the first five weeks their values decreased from 800 to 600 milliseconds and seven to four semitones respectively, the interval between six and twenty weeks is characterised by individual differences, and finally between twenty one and twenty four weeks all four infants evidenced an increase in range, variability and duration from four to six semitones, and 600 to 1400 milliseconds, respectively. The authors reported high positive correlations between range and variability, but they claim that duration and range are at least partially independent. Moreover, it was found that females show higher fundamental frequency characteristics than males. The increase in fundamental frequency range and standard deviation manifested after the fourth week of life suggests an increasing ability to control a broad range of fundamental frequencies (Oller, 1980). The increasing duration of non reflexive vocalisations indicates that infants systematically manipulate and practise different durations of vocalisations.

Keating and Buhr (1978) charted the development of fundamental frequency in a longitudinal study of six normal infants from eight months to three years. This particular corpus of data shows, similarly to Leufer and Horii (1977), that even in older infants' vocalisations, both very low and very high fundamental frequencies are common at all stages of language acquisition, the range of fundamental frequency obtained being from 30 to 3000Hz. Moreover, no developmental trend was evidenced in the overall fundamental frequency over this period.

Robb and Saxman (1985) considered the increasing linguistic ability as a potential influence on acoustic measures of speech production, and attempted to specify linguistic ability as a subjective characteristic. They obtained fundamental frequency values for fourteen children between the ages of eleven and twenty five months, an age period characterised by changes in physiological and linguistic development. The results showed that the overall mean fundamental frequency for the fourteen subjects was 357Hz, the overall mean fundamental frequency onset value was 363Hz and the overall mean segment duration was 357 milliseconds. Unlike Keating and Buhr (1978), this research demonstrated a decrease in average fundamental frequency of 86Hz with age (from 400Hz between 11-13 months to 314Hz between 23-25 months), as well as in variability and an increase in the mean length of the utterances. The decrease observed in average fundamental frequency as a function of age can be explained by the physical growth in children, that is the increases in vocal fold length and mass (Hirano et al,

1981). The decrease in fundamental frequency variability, though, is less easily explained by growth alone; it might involve an interaction between maturation of the nervous system and the gain of control over laryngeal adjustments for linguistic function (Netsel, 1981).

Delack and Fowlow's study (1978) is one of the very few that report findings on gender differences in prosodic characteristics. The mean fundamental frequency remains fairly stable at 355Hz over the first year of life, with females generally exhibiting a higher fundamental frequency throughout this period (20 to 25 Hz). The males, however, drop to 325Hz by the end of the first year. Differential trends by gender were also found with respect to duration and range. In general, there is a mean fifty percent increase in duration over the first year. For both genders, range increases by about twenty percent (from 85Hz) up to the age of six months and continues to rise to 110Hz for females, but drops back to 80Hz for males by the end of the first year. Regarding contour shape, the distribution is essentially identical for both males and females. The Rise-Fall pattern is not only predominant, but also exhibits the most notable changes with respect to age, increasing from forty to fifty five percent during the first year. The distributions of the other contours do not alter significantly over time.

One can note some diversity in the results of studies measuring acoustic features in infant vocalisations longitudinally, partly due to the different methods used for obtaining these measurements. Furthermore, one needs to be cautious in treating acoustic measures of speech production as indices of vocal tract maturation. While the infant's vocal tract is remodelled by about four months to resemble that of the adult, developmental changes continue in a number of respiratory, laryngeal, and supralaryngeal structures. The various tissues (muscle, bone, cartilage, and other tissues in different parts of the vocal tract) do not have a linear rate of growth, nor do all tissues grow at the same rates. This situation is typical of ontogeny (Kent and Miolo, 1995). Variations in development of vocal production of infants have been used to support a dynamic systems interpretation of behavioural development (Thelen, 1991). However, common features of prosody in infant and adult vocal productions give strong indication that morphogenetic constraints are applied in the developing brain over and above modifications due to peripheral developments in the vocal mechanisms (Trevarthen, 1973).

(b) The Linguistic Tradition in the Study of Infant Prosody

The linguistic tradition is interested in the emergence of various aspects of the adult linguistic system, and it searches for formal and functional continuities between characteristics apparent in infant vocalisations and those of their mother tongue.

In a longitudinal study of two infants, Stark and colleagues (1975) described in detail and compared three types of early vocal behaviour namely, crying, discomfort and vegetative sounds on the basis of breath direction, voicing, pitch, loudness and glottal and superglottal constriction. According to the results, not only do these reflexive sounds show 'suprasegmental' structure, but also some elements of this structure are incorporated in cooing sounds, which eventually merge into the acoustic productions of speech. Followers of this research stream (Lewis, 1936; Lenneberg, 1967; Crystal, 1986) have asserted that intonation is one of the earliest linguistic features acquired by the child, well before the development of recognisable words. Language-specific learned patterns of prosodic behaviour have been observed from around six months. Before this period, no difference has been observed between the vocalisation patterns of infants from different language backgrounds.

In a study of 'babbling' production of six, eight and ten month old infants from different language backgrounds, de Boysson-Bardies and colleagues (1984) found that adult judges were able to identify infants from their own linguistic community on the basis of voice quality and tonal contrasts present within long coherent intonation patterns. Furthermore, in late babbling, infants seem to increasingly gain control of the temporal organisation and the intonation of their language, especially the terminal contour, and adults can reliably identify and categorise such intonation contours as 'questions' and 'commentaries' (de Boysson-Bardies et al, 1982). It seems that certain 'suprasegmental' features of the target language are acquired very early and that the vocal tract becomes attuned to the laryngeal and supralaryngeal settings that are specific to the target language (de Boysson-Bardies et al, 1984).

Crystal (1979) suggests that, at the beginning of life, infants produce undifferentiated, biologically determined vocalisations, the so called 'basic cry'. In the next few months largely innately determined affective vocalisations are differentiated. At around seven months the 'Primitive Lexical Items' appear; these are the result of an imitative process of adult forms, which the infant perceives as units with a specific phonological shape and primarily non-segmental character. These patterns are formed

by the configuration of pitch, rhythm and pause and they are supposed to have a major grammatical function, i.e. the substitution of one pattern for another would lead to a different structural description of the utterance in terms of a grammatical model derived from the adult language. By producing 'Primitive Lexical Items' the infant develops an awareness of longer prosodic units with prosodic pattern similar to that of adult sentences. By eighteen months almost all the distinct intonational movements of the adult language have been used, though not all the adult meanings. According to this interpretation, linguistic development of utterances does not seem to begin with a composition of individual, independently movable items but as a whole tonal pattern with little form. With further development this whole becomes differentiated into component parts (Lenneberg, 1975; Weir, 1966).

A representative of research suggesting formal as well as functional similarities between prelinguistic and adult 'prosodic contours' is the study of Tonkova-Yampols'kaya (1969). A gradual approximation of 'comfort intonations' to those produced by adults was found, whereas 'discomfort intonations' are similar to the adult patterns from birth. In particular, between two and seven months an 'indifferent intonation' was discerned in infant vocalisations that was structurally comparable to the adult intonation of 'assertion' and 'enumeration'. During the sixth month the intonation of 'happiness' is differentiated. At seven months infants produce 'request' intonation and an expressive calm cooing, which corresponds to affirmation in adults. The number of such intonation types rose sharply during the ninth month. Finally, 'insistent' intonation in infant vocalisations corresponds to 'persuasion' and 'insistent command' in the adult language system and a 'question' pattern appears at thirteen months. The author underlines the point that the intonation patterns of infant sounds are not identical in structure to the adult ones, but they are to a considerable extent similar to them. The behavioural evidence is, however, inadequate to indicate such functional intentions in the child.

Menyuk and Bernholtz (1969) demonstrated that adults could classify children's one word utterances as 'declaratives', 'questions' or 'emphatics' with agreement more than 80%. Moreover, spectrographic analyses of the utterances yielded a typical fundamental frequency contour for each type of utterance.

Vihman and Miller (1988) provided data suggesting that the transition from 'babble', i.e. any vocalisations not identifiable as words, to words is a gradual one insofar as the phonetic shapes are concerned, that is, that the repertoire of sounds and sound combinations used by a child in late babbling and in contemporaneous early

words are closely related. According to the authors, words are crystallised slowly out of the pool of vocalisations which reflect the varying phonetic preferences of a child, by means of imitation of adult sound patterns. A typical example of a speech-like form around the end of the first year is the overlay of unequal stress and contrasting intonation patterns on long vocalisations.

It has been suggested that since children often impose interpretable intonation contours on single word utterances, these prosodic elements indicate underlying linguistic structure, i.e. they serve a syntactic function. For instance, when a child utters a rising terminal intonation contour with a single word, and adults respond as though he/she uttered a request, then, this is taken as an evidence that the child must possess some knowledge of the syntactic structure of interrogation (Lenneberg, 1967). A debate follows on whether linguistic syntax guides the learning of intonation or vice versa.

Bloom (1973) offers a counter-argument to the viewpoint that linguistic structure underlies the use of intonation patterns during the one word stage. In particular she observes that single-word utterances are not phonologically contrastive. That is, the child's intonation patterns cannot be systematically correlated with systematic intentions; a rising terminal contour, for example, does not always signal a question. Thus, the author suggests that the children do not possess any systematic knowledge of structure; rather, they learn to use the intonation pattern of the target language after they learn that basic grammatical distinctions are signalled by word order in English, that is after they learn syntax.

Although Weir (1962) does not directly deny that sentence melody is one of the earliest linguistic features in infant vocalisations, she asserts that this is true only as far as the purely imitative phonetic aspect of language acquisition is concerned. On the other hand, she observed that although her subject could utter four pitch levels and three types of pitch contours namely, falling, rising and level, these were not used contrastively as they are in standard English. Only the highest pitch level was related to calls and urgent requests, and it was classified as part of the emotive function of language. The author concludes that "to take an utterance and assign it a certain meaning on the basis of standard English intonation is most misleading" (Weir, 1962 p. 29).

In the same cautious vein, Weeks (1982) claims that although infants produce a wide range of intonation patterns in their babbling, sometimes even very complex ones,

they have not mastered the intonation system of the target language; "what they apparently do is to note that speakers of the language do not speak in a monotone" (Weeks, 1982 p. 159). In the second half of the first year what infants have acquired is the broad, largely universal prosodic patterns, but not the fine nuances of the adult language. Infants develop their own rule-ordered intonation system, just as they develop their own systems of phonology, morphology and syntax, none of which matches the adult system at this early stage. The intonational system of the adult language is only mastered after three years. However, the early idiosyncratic intonation system is meaningful to the infant and this meaning is usually communicated successfully.

Crystal himself (1973b) emphasises that the linguistic approach to infant prosody carries the risk of uncritically describing and categorising the features of infant vocalisations in terms originally defined for the study of adult language, or reading too much adult meaning into the infant's utterances. He observes that so far researchers have not been cautious enough when classifying infant phonological patterns and attributing meanings to infant vocalisations; this is even more disturbing if one takes into account the fact that intonation contours, which may have powerful interpersonal significance, do not have single meanings, but can be used in a variety of semantic contexts (Crystal, 1973b).

In any case one cannot describe a developing system with categories defined for the description of an already developed system. Reinforcing this argument, I would add that studying the role of infant prosodic patterns from an adult perspective means that infant vocalisations are seen simply as a reduced form of adult language. Thus, with this approach one cannot approach the origins and the nature of infant vocal development in its own right as part of an early communication system that is effective before language. As Studdert-Kennedy (1986) states, the linguists must avoid the temptation of seeing each of the infant's vocal utterances simply as a precursor to words in the native language. Further, although the linguistic approach illuminates many important and empirically documented trends in the early development of language, it does not provide any hypotheses about the underlying mechanisms which could account for language acquisition (Kaplan and Kaplan, 1971).

Oller (1980, 1986) approached the issue of <u>form</u> continuity between infant and adult vocalisations with an innovative method named '**Metaphonology**' which overcome satisfactorily, I think, the problem of describing infant vocalisations in terms of adult categories. Metaphonology is actually applicable to any communicative sound

system linguistic or non-linguistic. This technique specifies how general acoustic parameters manipulated in the world's languages, i.e., frequency, resonance, intensity and timing, are manipulated to generate well formed concrete phonological and phonetic units that is, vowels, consonants, tones, intonation contours and juncture cues.

Infant vocal development is estimated by the extent to which formations of these acoustic parameters are similar to the corresponding adult ones. According to the author, this method arose as a unifying model between two opposite methods which have been employed in the analysis of infant vocalisations, the instrumental acoustic and the phonetic transcriptional, but it is still distinct from both of them. Both these approaches have severe disadvantages, and neither has succeeded in detecting continuities between adult and infant speech (Oller, 1986).

An analysis of infant sounds on the basis of metaphonology revealed that at different stages in their phonological development infants master or practise different aspects of linguistic phonology. In particular, during the first month of life infants produce 'quasi resonant nuclei'; these constitute the introduction into the child's vocal repertoire of non reflexive vocalisations with normal phonation which is a basic feature of all mature languages. In the second and the third month by the production of COOs infants begin to control the contrast between closure and opening of the vocal tract during phonation. Closure and opening oppositions also constitute an important characteristic of languages. COOs consist of postalveolar (usually velar) consonant-like elements, often voiced fricatives, optionally combined with quasi-resonant nuclei. During the exploration stage between four and six months infants produce a variety of vocalisation types such as fully resonant nuclei with which the vocal tract's capacity for resonance is explored, 'squealings' and 'growlings' that represent manipulation of the vocal pitch, and yellings that represent manipulation of amplitude.

The period between seven and ten months is named the 'Canonical Babbling Stage' and is characterised by the production of sequences of units formed by consonant and vowel like sounds in a timing relationship that conforms to mature language restrictions. During the variegated babbling stage (11-12 months) relatively rigid syllabic characteristics of languages continue to appear in many of the infant's utterances but now infants systematically produce utterances with differing consonantal or vocalic elements. In addition, during this stage infants often produce a category referred to as 'gibberish'; these are phonetic sequences with contrasts of syllabic stress. The 'gibberish' and the variegated babbling sounds demonstrate considerable control over fundamental frequency contour which has begun in the first month of life.

Taking also into account the results reported by Leufer and Horii, Oller (1980) concludes that infant sounds represent explorations of the capacity for speech. Stark (1980) reports the same stages of speech development in the first year of life and notes that it is not clear at exactly what point the infant begins to be influenced in its output by the linguistic environment.

A number of linguists seem to have realised the problems inherent in trying to detect functional continuities between infant and adult vocalisations, and in their work one can find the germs of a more fruitful approach which strives towards establishing the role of infant prosodic patterns in expressing the infant's feelings or intentions in particular situations.

In his classical detailed account of language acquisition relying on impressionistic observations Lewis (1936) noted that as early as in the first month an infant will normally be using different 'intonational' contours, innate in origin, to express different affective states. Moreover, these contours were claimed to express intention, since this psychological element enters into the infant's play within the first three months of life, but they are not objectively referential. In the course of the next few months, further differentiation of intonation takes place and different intonational patterns have a 'declarative' or a 'manipulative' function that is, they express emotions or they direct the other's attention to the topic of the infant's interest. The sounds gain increasing reference to the situation. The author asserts that this achievement is the result of the motive to bring others within one's circle of activity and to share feelings and knowledge. The vocalisations produced during this period carry the rudiments of reference.

Werner and Kaplan (1963) offer a thorough exploration of the functions of prosodic patterns in prelinguistic communication in relation to the development of reference. According to these authors infants for the most part of the first year of life tend to imitate all kinds of noises. From nine months these intonation patterns are combined with gestures to express attitudes of contact or demand where the referent is identical to the addressor's needs. At this point the various attitudes are more or less syncretically fused, although one or another might be dominant in a particular expression. Progressively, the intonation patterns together with gestures will come to serve primitive referential functions as in the case of declaratives, where the object of discourse is relatively distinct from the addresser's needs. At the end of the first year

articulated entities will start to take over the referential function whilst the intonational patterns will be expressing the addresser's attitudes.

According to Carter (1978), the infant's preverbal idiosyncratic vocalisations are systematic, in the sense that they are highly correlated with intentions and/or situations. Nevertheless, the author points out that, since one cannot adequately access the extent to which the infant's reality matches that of the adult, and thus cannot know what features, real or imagined, of the situation or the infant's internal state elicit a given vocalisation, one is constrained from reliably assigning a meaning to that utterance.

Following the theoretical viewpoint that intonation is among the earliest associations of language form with aspects of meaning, Flax and her colleagues (1991) examined the hypothesis that infant prosodic patterns would continue to serve the same functions in the prelinguistic and during the single word period, until the child has other conventional means of conveying the same communicative intent. Three children were observed interacting with their mothers at three different times: before the onset of single words, when vocabulary consisted of ten words, and when vocabulary consisted of fifty words (age range 11-22 months). Although differences were found between them, the children were similar in that the greatest proportion of their rises were produced in vocalisations that were coded as forms of requests, whilst the non-rises were produced in vocalisations codified as 'Interactive Comment', 'Response', 'Non-Interactive Comment', 'Command'. Moreover, categories with emotional content such as 'Protests' and 'Request for Attention' tended to have higher centre and peak fundamental frequency than those less emotional expressions such as 'Responses' and 'Non-Interactive Comments'. These relations between communicative function and terminal contour and communicative function and peak or centre fundamental frequency remained relatively constant for each child, from the prelinguistic period sampled through the single word period. Adults use the rising contour type in the same way.

This study can be criticised for the method of prosodic analysis employed (see Chapter 5) as well as in regard to the coding system for assigning a communicative function to each vocalisation. The definitions of the communicative function categories are impressionistic, and not relying on the accompanying non-vocal behaviour of the two partners; thus, the definitions remain imprecise. Commenting on their results, the authors (Flax et al, 1991) note that since the associations between rise and non-rise contours and communicative functions demonstrated in this study were based on relative distributions and not necessarily on frequencies of occurrence of each category,

it is difficult for one to conclude that the children had induced any relationship between prosodic variables and the communication of a particular function.

This study starts off with a linguistic hypothesis, which postulates that certain prosodic variables are possibly associated with aspects of social interaction before they are used linguistically. Then, the authors attribute these associations to the vocal input the children observed to be related with particular interactive roles, or to a learning process which takes place whenever adults interpret chance rising contours that are produced by infants as evidence of requests and reward them, thereby teaching the children that rise is to be used for requests. This interpretation of the results ignores the active role of the infant in the communication with the caregiver, and the contribution of the infants own expressive potential- an important part of which is prosody-to the transmission of messages about his or her feelings and interests.

Marwick (1986) approaches the issue from a different angle. Instead of searching for any syntactic structure in the infants' intonation contours, she investigated how the mother's intonation is related to contrasts in certain communicative functions in relation to the child's intonational patterns, as well as the mother's response to her child's intonation patterns. Two infants, aged from sixteen to twenty eight months, were studied with their mothers. The findings demonstrated that both mothers used intonation as one element in an integrated expressive code to clearly convey their communicative intentions during the interaction with their child. As the messages to be conveyed became more complicated, so did the intonation contrasts. The children's intonation closely matched that of the mothers in form and function. The children also clearly did not use intonation contrastively, in the sense of one intonation pattern or group of patterns indicating only or always one function, and another intonation pattern or group of patterns a different function. Rather, it seems that the child's intonation should be viewed as part of a package of elements which combine to convey the communicative function of an utterance. Furthermore, the mothers responded to their children's intonation according to the ways in which the mothers themselves use intonation and not differentially in terms of intonation pattern. This fact suggests that the mothers do not impose any kind of simplistic rising versus falling intonation organisation upon the child's utterances, but rather they exert an overall influence upon the child's communicative development.

(c) The Application of Speech-Act Theory in the Study of Infant Prosody.

Jerome Bruner (1975) and John Dore (1975) recognised, more explicitly, the need to trace the origins of language acquisition in the communicative functions served by prelinguistic vocalisations, and in their early work they based this working hypothesis on the Speech-Act Theory.

The 'Speech - Act Theory' proposed by Austin (1962) and extended by Searle (1969) derived from the Wittgensteinean idea that there is functional diversity in the utterances of natural languages and that the meaning of a word is revealed in its use (Wittgenstein, 1953). In his famous treatise 'How to Do Things with Words' (1962) Austin claims that when one produces an utterance one is actually performing a communicative act, in order to accomplish a goal. A speech-act consists of three elements (a) the locutionary act which is the production of an utterance with conventional form and reference (b) an illocutionary act which conveys how the speaker intends his utterance to be understood, and (c) the perlocutionary act which conveys the actual effect an utterance had on the hearer. Austin pointed out in his discussion of speech-acts that there are various 'felicity conditions' which an illocutionary act must fulfil in order to be successful and non-defective. These felicity conditions concern the appropriateness of the utterance to the circumstances in question, the sincerity of the speaker, the consistency with what is entailed and implied, whether or not the speaker is in a position to state what is being stated and the nullity or otherwise of presuppositions.

Applying the 'Speech-Act Theory' to the problem of language acquisition, Dore (1975) phrases the question of how the infant acquires the linguistic conventions necessary to express its intentions; in other words, how speech acts develop. He revised the notion of speech act and described the single word utterances of one year old infants in terms of 'Primitive Speech Acts'. The 'Primitive Speech Act' was defined as containing a rudimentary reference expression (the word) and a primitive force indicating device (typically an intonation pattern). It is not merely an elliptical adult speech act, but a qualitatively different entity which possesses only some features similar to adult speech acts. This approach assumes that at the single word stage the child has acquired at least those linguistic devices necessary to utter a single concept and that prosodic patterns accompanying the production of single words indicate that the child is able to communicate its intention with regard to that notion. Further, a speech-act approach utilises contextual features in a theoretically satisfying way. Such

features are considered as clues to the child's intentions, but not as part of its linguistic knowledge *per se* .

In his early work Bruner (1975) also adopts the 'Speech Act Theory', although with a broader perspective than Dore, in order to consider language development and specially the transition from prelinguistic to linguistic communication in relation to behaviour generally, as well as to allow for an emphasis on the use of language rather than its form. The author claims that what may be innate about language acquisition is not linguistic in nature, but some special features of human action and attention that permit language to be decoded by the uses to which it is put. This capacity leads the infant to first learn prelinguistically to make the conceptual distinctions embodied in case grammar. In the same vein, it is suggested that infant prosody is an imitatively learned phonological pattern, which constitutes a prosodic envelope or matrix into which the child 'knows' that morphemes go, and which may have an interrogative, a vocative/demand or an indicative contour. The conceptual distinctions as well as the learned prosodic patterns are developed in a mutual relation between the infant and a speaker of the language and aid the acquisition of 'Speech Acts'. Bruner himself considers the process of learning prosodic patterns as puzzling and claiming lack of evidence he does not adequately explain the role of the infant prosodic patterns in language development.

Following the Speech-Act Theory, Linda Ferrier (1985) studied the role of rising pattern in the vocalisations of twelve-month-old infants' in relation to the rising intonation patterns produced by their mothers during communicative episodes. The author claims that intonation, along with other paralinguistic communication systems, is one of the early means by which infants communicate with their caregivers. The participants were 10 infants being brought up in English-speaking families, and their mothers. Each mother and infant utterance was entered in one discourse class and was also attributed a speech-act function. The results showed that infants' use of rise is equally distributed between the two eliciting discourse positions of 'initiate' and 'continue'. Moreover, all the infants in the group used rising intonation with approximately the same proportional frequency as their mothers and generally to perform the same set of speech-act functions. Within this broad picture of the impact of maternal input, variance was observed which was associated with individual strategies. The author notes that these results do not suggest intonation use as the one and only route to early language acquisition. However, it is perhaps surprising that infants use intonation so systematically at a stage when many of them are producing only one or even no recognisable adult words (Ferrier, 1985).

The 'Speech-Act Theory' gives explicit recognition to the interpersonal dimension of language by claiming that producing an utterance is essentially engaging in a certain kind of social interaction. The theory is not restricted to communication by means of spoken language. Indeed, it is arguable that there are certain non-linguistic communicative acts that would satisfy the definition of 'speech-acts' (Lyons, 1977). However, Austin developed his theory with particular reference to a linguistic system. The speech-acts are "in general made possible by and performed in accordance with certain rules for the use of linguistic elements" (Austin, 1962 p. 16). Thus, an analysis of illocutionary acts must capture both the intentional and conventional aspects and especially the relationship between them.

(d) Cognitive / Rationalistic and Social / Pragmatic Approaches to Infant Prosody

The Papouseks (1981) examined in detail the 'musical elements' in their daughter's vocalisations produced during interactional episodes or when the infant was alone, from birth to sixteen months. Tanya's initial crying showed considerable variation in prosodic patterns. From the first month onwards, musical sounds were produced with certain pitch, timbre, loudness and rhythm characteristics. The musical quality of these tones improved with age and towards the end of the third month they became so differentiated as to allow the infant to exhibit a rather rich variety of tones in relation to the situation. At the age of four months this infant seemed to explore the capacity of the vocal tract for the manipulation of pitch and amplitude. This finding is in agreement with what is reported by Oller (1980). The pitch range varied considerably with the types of vocalisations, but it also seemed to increase with age covering almost three and a half octaves between ten and twelve months.

During the period of reduplicated babbling a variety of melodic patterns appear and two parallel lines become obvious namely, the speech-like intonation contours accompanying the linguistic differentiation, and the musical melodic contours in imitated or invented songs. Speech-like contours first appear in the infant's monologues as practising. According to the authors these prosodic patterns not only reflect distinct syntactic functions involved in the adult language as some linguists have suggested (Lenneberg, 1967), but also closely mirror the features of baby talk addressed to her, which convey biologically relevant and physiologically effective messages.

Although the Papouseks consider infant vocalisations as one of the most important features of the communicative exchanges between parents and infants, the

fundamental function they attribute to them, and consequently to their prosodic features, is purely cognitive. In particular, they claim that infant prosodic expressions contribute to cognitive-integrative operations involved either in social circumstances, or in the infant's isolated play.

Certain vocal cues have been found to accompany processing of information, reaction to moderately discrepant stimulation (Zelazo et al, 1975) anticipation of familiar events, fulfilment of expectations, or evaluation. Especially in a communicative context, the signalling of fluctuating states of comfort/discomfort through 'musical elements' functions as a feedback response to preceding parental stimulation, offering information on the infant's readiness to attend, to interact, and to integrate interactional experience. Microanalytic studies of naturalistic parent infant interactions have demonstrated that parents use infant vocal signals for adjusting the amount and quality of stimulation to the infant's momentary readiness to interact and to integrate experience and for responding contingently with adequate forms of intuitive didactic caregiving behaviours. This contingency of parental responsiveness on infant vocalisations provides the infant with abundant experience that vocalisations may elicit predictable behavioural responses. Such experience may enhance the development of intentional communication, an important precursor of speech acquisition (Papousek and Papousek, 1981; Papousek et al, 1984; M. Papousek, 1989; M.Papousek, 1992).

However, mother-infant interaction cannot be described simply as a stimulus-response process of information exchange or management of cognitive processes. Dialogues between the mother and her infant rely on mutual 'attunement' of feelings and interests (Stern et al, 1985) generated by innate motives for communication. They are clearly related to intentions, not only the perception and production of events.

A rationalistic approach to language acquisition supports the idea that language is generated *in* the infant's mind, as a product of structured thought. Piaget (1951) sees language as simply as a by product of the general cognitive development, providing an economical means of representing reality, and he describes language development within the framework of the transition from sensory-motor to representational modes of thought. Piaget claims that language is an important factor in both the formation and the socialisation of representations Further, he notes that the non-verbal roots of representation are to be found in imitation and play, which have also been considered as cognitive phenomena. Bates and colleagues (1976) have demonstrated that play, imitation and means-end relations are good predictors of gestural and verbal communicative development. The Piagetian perspective emphasises the cognitive

achievements of the child as a crucial prerequisite of language, thus, attributing an 'egocentric' character to language acquisition, and neglecting the role of intersubjective exchanges. Ultimately language development is seen as dependent on the progress of the child's own thought.

However, recent research has demonstrated that imitation-which is the basis of representations- is an intersubjective phenomenon more than anything and its role is to regulate communication between the infant and the caregiver (Kugiumutzakis, 1993); the same is the case for play (Reddy, 1993).

A research stream on language development that follows the Piagetian tradition regards the study of 'Pragmatics' that is, the description of rules for relating an utterance which is the product of a combination of linguistic, cognitive and social rules, to a context, where meanings can be found. Pragmatics occupies the interface between linguistic, cognitive and social development (Bates et al, 1975). The authors note that it is not appropriate for psychologists to apply linguistic concepts directly without considering their fit to current knowledge about cognitive processing and social dynamics. The pragmatic approach, drawn on principles of Speech-Act theory, studies three aspects involved in speech:
(a) PERFORMATIVES- which describe the capacity to formulate and execute a communicative intention
(b) PRESUPPOSITIONS- assumptions about the context that are necessary to make that sentence verifiable, or appropriate or both, and
(c) CONVERSATIONAL POSTULATES- a particular class of presuppositions about the nature of human dialogue in general.

Applying the pragmatics approach to language development Bates and her colleagues (1975) claim that since performatives essentially reflect the intention to communicate, the earlier possible appearance of performatives requires the emergence of intentionality which is the result of certain cognitive developments of the Piagetian sensorimotor stage 5. In particular, it has been suggested that intentionality begins when the infant becomes capable of differentiating ends from means. The two performative structures that seem the most basic to prelinguistic and linguistic communication are the 'imperative' and the 'declarative'. These perfomatives are expressed with vocal and gestural signals or combinations of these modalities.

Dore and his colleagues (1976) also followed the Piagetian approach. They videotaped four children (three boys and one girl), all from middle-class backgrounds,

once a month from 11 to 16 months. The child's mother and a nursery school teacher were with the child during the sessions which centred on free play in a familiar setting. In the early stages of the investigation it was observed that the children produced Phonetically Consistent Forms (PCFs) (Dore, et al, 1976) or 'Indexical Expressions' as he later calls them (Dore,1983), which are transitional between prelinguistic babbling and words. In terms of form these vocalisations have four main characteristics:

(1) are readily isolable units, bounded by pauses (unlike babbling),

(2) are more phonetically stable than babbling though less so than words, and have distinctive prosody

(3) can be loosely correlated with features of the environment and/or the infant's behaviour. Thus, their production is neither random as in babbling nor as rule-governed as words, and

(4) occur repeatedly as items in a child's vocal repertoire.

An analysis of the functions of the PCFs revealed four distinct types:

(1) Affect expressions: These are expressions of mood and attitude,

(2) Instrumental expressions accompany acts of striving and involve a directed gaze toward an object or adult with the apparent intention of obtaining the object or engaging the adult in interaction,

(3) Indicating expressions are communicative but not directive; most often they involve pointing but not dissatisfaction when the action fails to ensue, and

(4) Grouping expressions reflect an interaction between subjective state and attention to objective properties.

There are two types and two phases of development of indexical expressions. The first, affect indexicals, seem to emerge from early crying and delight sounds, and at the beginning are exemplified primarily by grunts of protest or squeals of pleasure. They are occasionally directed to someone via gaze, and are dominated by gestures, facial expressions and tones typical of aroused affect states. In such behaviour the affect is more salient than the form is conventional, and these indexicals tend to be replaces by forms like 'stop' and 'no' later in development. The second type, formal indexicals, exhibit less affect indicating behaviours, and increasingly approximate the surface forms of adult speech. They are better defined syllabically, phonemically more stable, and more varied prosodically than the affect type.

Speculating on the nature of these vocalisations Dore (1983) characterises them as intentional expressions. He asserts that the transitional period between babbling and the first words at the end of the first year is in between two other kinds of intentionality: the

intent to achieve a physical goal and the intention to convey a word meaning. Intentional expressions are more than physical actions, but less than intended word meanings. This period itself can be subdivided to two phases: 'intending-*in*-expressing' and 'intending-*to*-express'. Vocalisations of the first type are closer to intent-to-act, insofar as they accompany actions and are not addressed directly to people. They express a state but are not intended to express a specific state. They are not mediated, premeditated, or planned, but rather immediate, spontaneous, and reflexive. In contrast, the 'intending-to-express' vocalisations are closer to the intention to convey word meanings. They do not merely add more of the same to an action but rather supplement it usually involving something to be done or said about the object. They are directly addressed to someone, with sharp focus of personal gaze to indicate the other's status as addressee. They are premeditated and they are 'representational' intentions to the extent that some kind of deliberation takes place, some rudimentary cognitive content is executed.

Although Bates and her colleagues (1975) and Dore (1975) recognise the importance of communication in language development, they rely on the assumption that intention is the deliberate pursuit of a goal by means of instrumental behaviours subordinated to the goal. So, intentional expression requires the establishment of means-end relation, they believe in accordance with Piaget, which only happens the second half of the first year. Nevertheless, the philosopher Brentano (Brentano cited in Hubley, 1983) stated that all kinds of psychological action namely, perception, learning, cognition, emotion and communication are intentional since they have reference to or are directed to other persons, objects or events.

In the first two months infants already show evidence of perception (Eimas, 1974), learning (Papousek, 1969), imitation (Kugiumutzakis, 1985), emotion and communication (Murray, 1980). As Hubley (1983) notes it does not seem to be a fundamental change in the infant's nature from that of a non-intentional being to that of an intentional one, but changes in the complexity and specificity of intentional function and expression.

(e) Studies focusing on the Communicative Functions of Infant Prosodic Patterns

A theoretical viewpoint on language acquisition that would appear to be more appropriate for studying infant vocalisations is one that proposes the origins of language to be searched for in the dialogue *between* the mother and her prelinguistic infant, i.e. in their communicative relationship and its regulations (Kaplan and Kaplan, 1971; Ryan,

1973; Bruner, 1975, 1978, 1983; Halliday, 1975, 1978; Dore, 1983; Locke, 1993). Some authors concerned with the means of adult language adult language have emphasised that intersubjective communication is the essence of any a linguistic exchange. Habermas (1970) asserts that adult verbal communication in society cannot be understood only as an application of linguistic competence, but rather has the structure of intersubjectivity that makes such application possible also to be explained.

The main idea underlying the communication or 'socio-linguistic' theories of language development is that the form of human semantics is to be explained as a consequence of adaptive processes in language, which has evolved as the main expressive means of communicative functions and feelings (de Laguna, 1963; Halliday, 1973; Locke, 1995). As Locke notes "By saying words, speakers provide themselves with opportunities to *give voice* to these feelings. For speaking permits individuals to indulge in vocal displays and to do emotionally expressive things with their eyes and mouth" (Locke, 1994 p. 13). The transmission of emotions and communicative functions through language brings about human cooperation; in other words, it coordinates the diverse activities of humans for the attainment of reciprocal goals and thus, facilitates cultural creativity. The course of language development must be sought in the light of this function (de Laguna, 1963). Halliday (1973) suggests that language has many uses which have sprung from four fundamental functions: (a) the Instrumental function refers to the use of language as a means of getting things done, (b) the Regulatory regards the use of language to regulate the behaviour of others, (c) the Interactional refers to the use of language in the interaction between the self and others, and (d) the Personal refers to the awareness of language as a form of one's individuality.

Ontogenetically speaking, human beings do not learn the functions described above (Halliday, 1973). Rather, they are endowed with the capacity to express these functions in communicative exchanges with an empathic caregiver. The dialogue between the mother and the infant is mutually regulated by both partners through a continuous negotiatory process. Although the infant and the mother express themselves in different forms, since the mother also uses conventional expressive forms as an active member of a linguistic community, they ultimately manage to imitate, match, complement, and even mismatch, each other's motivational state, keeping the flow of transmission of feelings and interests going in a highly dynamic and mutually regulated process.

A number of studies on early language acquisition exemplify the socio-linguistic / intersubjective approach to early language.

Halliday (1975) gave a thorough account of the functional development and the vocal patterns of his own son's vocalisations. As he observed, at nine months it was possible to identify just two expressions which apparently fulfilled the criteria for being language, that is they had constant meanings interpreted in terms of functions: the *Interactional* which reflected the motive for companionship ("let's be together") and was conveyed by a mid or mid-low falling tone, and the *Personal* which reflected interest in the movement of a specific object; this function was also expressed by a mid or mid low falling tone but usually of wider range. Nigel also had another three meanings expressed gesturally: two *Instrumental* meanings reflecting the infant's will to have or not to have an object and one *Regulatory* meaning by which he was directing the adult's action on an object. These gestures disappeared at around ten months.

By ten and a half months Nigel had expanded his functional repertoire. At this time he was using a mid falling tone to convey the *Instrumental* function ("give me..."). In the *Regulatory* function two meanings could now be distinguished which were conveyed by distinct prosodic patterns: a) A Command directed to a specific individual, requiring one to repeat something that had been done immediately before. This function was also expressed by a mid falling tone. b) An intensified form of the first one which was conveyed by a falling tone of wider range. When Nigel initiated *Interpersonal* contact ('greeting') his vocalisations had mid pitch, narrow range and falling contour. These utterances were also used to direct the adult's attention to an object as a channel for interacting with the other person. If the invitation was intensified, the pitch range was wider; this form was an impatient greeting which was not mediated by any joint action. On the other hand, responses to vocal invitations for interpersonal exchange were low pitched falls. Finally, there was a set of meanings within the *Personal* function one of which expressed a state of withdrawal with a narrow ranged low pitched falling tone, and the rest a state of participation, involving the expression of some form of pleasure or interest. These meanings were conveyed by low pitched fall vocalisations. Nigel's system developed further and at nineteen months the child used a rising tone on utterances requiring a response and a falling tone on utterances for which no response was required. This Halliday interpreted as his son's own way of distinguishing between two broad types of language use, the 'pragmatic' and the 'mathetic'. The prosodic expressions described so far are not imitations of any

intonation patterns of the English language, rather they are spontaneous creations of the glossogenic process.

A criticism of Halliday's categories is that although the distinctions may be theoretically sound, he has not developed a precise and detailed coding system by which one could reliably categorise infant vocalisations.

Bruner (1982, 1983) focused on the development of 'Request' in what he calls 'formats'. The formats are play routines developed and facilitated through communicative exchanges between the infant and the mother. He distinguished three main types of request in two children and gave an impressionistic and simplistic description of their prosodic pattern. The first and the simplest procedurally, is *request for object* which is present at eight months. The second type is *invitation to an adult* to share a role in a game; such requests occur in contextualised, highly familiar routines and appear after thirteen months. A third type is *request for supportive action* which also becomes common after thirteen months. Both children produced stylised request calls for requesting objects, consisting of a particular phonemic form with a particular prosodic pattern, which, however, had rising end for one child and falling for the other. Contextual features such as insistence on the part of the child were observed to alter the overall shape of the contour. By the second half of the second year these vocal expressions were replaced by more speech like forms, that is sentence-like babble strings. Requests for invitation to joint action are accompanied by vocalisations that are notably less insistent than those accompanying object requests or requests for supportive action. As was the case with object requests, with increasing age the babbling accompanying invitations became longer and sounded more sentence like in intonation.

Following Halliday's tradition, Marcos (1987) investigated the communicative functions of pitch direction and pitch range during the second year. Two cohorts of infants were observed over an 18 week period. Group A was composed of six subjects, and was observed at 14, 15.5, 17 and 18.5 months. Group B was composed of four subjects and was observed at 17, 18.5, 20 and 21.5 months. This study demonstrated that initial pitch range is shown to be consistently higher for repeated requests for objects than for initial requests and for initial requests for objects than for labelling. Pitch range for giving and showing is often lower than for requests and higher than for labelling. Regarding the pitch direction parameter, in the youngest subjects no difference was observed for pitch direction for 'requests' vs 'labelling'. At 15.5 months though the proportion of rising tones tended to be higher for requests, and the

proportion of falling tones higher for labelling. This trend disappeared for the next oldest age group but reappeared in a more stable fashion at 18.5 months. However, this research has methodological deficiencies, since it is not explained how the different functions were attributed to infant vocalisations.

Delack and Fowlow (1978) studied the development of prosodic contrastivity in different contexts during the first year of life. Nineteen first born infants (seven female and twelve male) brought up in monolingual English home environments were filmed every fortnight from one month to approximately one year. The corpus of analysis consisted of vocalisations produced in four different contexts defined as social situations: (a) when the infants were alone (S) (b) when 'conversing' with the mother (SM) (c) or with another adult (SOP), and (d) when alone in the presence of various objects. The latter contexts were further categorised according to the sensory modalities namely, visual (SV), auditory and visual (SAV), tactile and visual (STV), or auditory and tactile and visual (SATV) by which the objects would be primarily, or at least most likely, apprehended.

The results demonstrate that each context may be described by a unique constellation of contours. In particular, the contours exhibited in the S context are complementary to the ones of the SM context; the first being mainly Fall-Rise or Rise and the latter Rise-Fall. The vocalisations produced in the S and SATV contexts show virtually the same pattern, although in the latter falls and levels were also found . The SOP context was characterised by Fall-Rise-Fall and Fall contours. In all the other contexts where the infant is alone with a variety of objects, falling contours predominate. The authors (Delack and Fowlow, 1978) conclude that these findings document the infants' capacity for vocal differentiation of the social environment events, which is one of the basic components in the development of communicative competence.

Although this approach is useful as one of the first systematic attempts to steer our interest to the role of infant prosody in prelinguistic communication, it is still inadequate. Considering prosodic differentiation as a result of the external social context implies a stimulus-response approach. However, prosody- as any other expressive means- conveys much more than a simple reaction to surroundings; rather, it reflects the infant's interpreted 'Psychological Context' (Ogden and Richards, 1923). This psychological context consists of a complex network of relations formed by the infant's apprehension of the present physical context, its understanding and interpretation of psychological aspects of the partner's behaviour as well as its previous

experiences of both these and the way they are linked with the present circumstances, all in relation to the infant's changing motives, emotions, interests and purposes. Thus, in a study of the role of prosody in communication, the concurrent non vocal behaviour should be described by a system which takes into account all the above parameters.

In a crossectional study of twelve children, six boys and six girls aged 23 to 25 months, Furrow (1984) examined the relationship between prosodic variables in children's spontaneous speech and aspects of their social behaviour while speaking. Three social context categories were identified:
(a) Eye Contact: Eye contact occurred during the expression of an utterance or in the preceding or following two seconds,
(b) Other Social: No eye contact occurred during the interval circumscribed by the two seconds preceding and following an utterance, but the child's behaviour involved an adult through physical contact or approach, or the adult's behaviour involved the child, or the utterance was near other utterances with social signs, or eye contact has occurred immediately before or after the two seconds interval, and
(c) Private Speech: No social marking accompanied an utterance or this was isolated from other utterances with social markings.

The results of this study demonstrate that children at two years of age vary prosodic aspects of utterances dependent on the communicative context in which utterances are produced. In particular, vocalisations produced in situations requesting a response i.e., during Eye Contact or Other Social contact are on average louder, and higher and more variably pitched than vocalisations produced in the other situations. On the other hand, vocalisations produced when children exhibit no sign of interpersonal behaviour are on average quietest and lowest and least variably pitched. Thus, the results suggest that children selectively control certain prosodic aspects of their speech for certain interpersonal purposes.

D'Odorico and colleagues (1991) made an intensive longitudinal study in a laboratory environment which attempted to overcome the problems of context definition stated above. Specifically, they investigated the hypothesis that prosodic patterns are selectively uttered in relationship to their production context by studying four infants aged from four to eleven months. The researchers state that the contexts were identified by a detailed analysis of communication occurring in infant-adult-toy interactions on the basis of discriminating dimensions relevant to the development of the capacity to take into account both objects and the psychological potentialities of the communicative partner simultaneously. This is one of the most important

communicative achievements in the first year of life. Most of the vocalisations recorded were produced in four contexts which were identified as follows: (1) 'Vocalisations during Infant Manipulation': this context includes sounds produced during or immediately after active manipulation or inspection of the toy. There are no glances at the adult either during or in the first two seconds immediately following sound production.
(2) 'Vocalisations during Shared Experience': this context also includes sounds produced during or immediately after active manipulation or inspection of the toy, but the infant looks at the adult either during or within two seconds of the vocalisation
(3) 'Vocalisations during Adult Manipulation': this context includes sounds produced when the adult is holding a toy and performing some action with it (either just showing it to the infant or manipulating it). The infant looks at the toy or the adult, or it switches attention rapidly from one to the other, and
(4) 'Vocalisations during Exchanges': This context includes sounds produced looking at the adult when neither of the partners is performing any action with the toy. The toy may be physically handled either by the infant or the adult but it is not considered by either. The results of the study demonstrated that vocalisations with different prosodic features are produced in different communicative contexts. It was found that 'Vocalisations during Shared Experience' and 'Vocalisations during Adult Manipulation' were more high pitched and had mainly rising contours compared to 'Vocalisations during Infant Manipulation' which have lower pitch and more level contours. The authors emphasise that at these early stages the sound-meaning system is idiosyncratic, and the relevance of specific features is different amongst infants. After about nine months the acoustic parameters used in this study no longer discriminate between vocalisations produced in different contexts. On the contrary a homogenisation process of acoustic patterning across contextual categories is apparent; it seems that a reorganisation process is going on both within and between the categories.

Despite the fact that the coding systems of Furrow (1984) and D'Odorico and colleagues (1991) make a step towards the definition of the vocalisation context in terms of the infant's psychological state at the time, they are not thorough enough to capture subtle characteristics of the situation that would reveal the infant's intentions to which a particular vocalisation is related. Emphasis is given to the gaze whilst other behaviours are neglected. What expresses the infant's motivational or emotional state is not simply the form of a behaviour but its expressive function; this function can only be identified if the form of a behaviour is considered in relation to the effects or purposes of the partner's previous behaviour. These crucial considerations regarding the

regulation of intersubjective contact were not taken into account by the authors in devising those coding systems.

One consistent finding deserves recognition. The results of studies adopting different approaches to context definition, and to prosodic analysis and including subjects from different linguistic environments (English and Italian), are consistent in showing that vocalisations demanding some intervention by the partner are likely to be marked by rising melodic patterns and high pitch (Halliday, 1975; Furrow, 1984; D'Odorico et al, 1991). This clearly relates to the infant's intent in cooperative awareness and the potentialities of the partner for facilitating the infant's purposes.

Shimura and her colleagues (1992) were the first to carry out a study on the expression of emotions by infants in the second half of the first year and the first half of the second. The researchers analysed through perceptual/acoustical experiments the vocalisations produced by a Japanese male infant during interactions with his mother over the period of 6 to 17 months after birth, to examine how adult subject perceive emotional profile of infant vocalisations. Perceptual judgements were made for 12 terms describing a variety of emotions by using a 5-point successive category scale. Three factors were extracted from a factor analysis which was carried out for the ratings. Factor 1 highly correlated with terms conveying positive emotions, Factor 2 correlated with terms describing negative emotions, and Factor 3 represented the 'Sad' emotion. 'Requiring' correlated with Factors 1 and 2. For the voice samples at 6 months, the average rating score was high for terms expressing positive emotions, whereas it was relatively low for the negative terms. At 17 months, the variation in ratings between the voice samples became large, and some voice samples were rated high also for negative terms. Moreover, preliminary acoustic analysis showed that voice samples assigned high 'Speaking' scores tended to have a falling fundamental frequency pattern. Voiced samples assigned high 'Requesting Affection' scores had mainly slow rising and slow falling fundamental frequency patters, while voiced samples with high scores for terms conveying negative emotions showed abrupt changes in fundamental frequency. These results indicate that even form 6 months infant vocalisations contain a wide variety of acoustic characteristics which may convey various emotional expressions. These acoustic characteristics are further enriched until about nine months; over the period of six to nine months especially develops the ability to express negative emotion without crying. The authors conclude that not all aspects of emotional expression through voice seemed to develop synchronously or simultaneously.

2.3 Are the Prosodic Patterns in Infant Vocalisations Meaningful ?

The discussion so far raises the pivotal issue of whether the prosodic patterns in infant vocalisations are 'meaningful'. Wolff (1969) was the first to consider infant cry vocalisations as having an inherent meaning. He distinguished four main types of infant cry vocalisations on the basis of systematic measurements on rhythmic and fundamental frequency features, and characterised them as 'hunger' cry, 'angry' cry, 'pain' cry and 'frustration' cry. These terms, in fact, describe what the author judged as the cause of cry. Crystal (1973b) objects that one cannot assume that the meaning of an infant vocalisation is what an adult sees as the obvious causative factor. He doubts the inherent meaningfulness of infant vocalisations. I would give a positive answer to this question; the arguments in favour of this are (A) the fact that the mothers spontaneously and often unconsciously attribute meanings to the utterances of their infants, and (B) because these prosodic patterns constitute part of a coherent expressive system which is functional the beginning of life and reflects the infant's motivational and conscious state at a given period. These two sets of arguments will now be considered in turn:

(A) Grice (1957) claims that one of the crucial prerequisites for a linguistic utterance to have meaning is to be interpreted as such by the hearer. Mothers tend to impute meanings to the vocalisations, gestures and postures of their babies from the beginning of life. Moreover, there is often a strikingly moralistic approach to these imputations (Bruner, 1975). The mothers of two month olds often reply to their babies as if they were telling a story (Murray and Trevarthen, 1986). Both parents can describe the vocalisations of their four month olds and attribute meaning and long term functions to them (Wong and Miller, 1984). Analysis of the linguistic content of mothers' spontaneous responses to infants during their communicative exchanges reveals maternal references to the infant's concurrent activities, state, needs or discrete emotions (Papousek and Papousek, 1994).

Keller and Schölmerich (1987) found that infants perform different types of vocalisations that can be interpreted as affective states from two weeks. These vocalisation patterns develop over the first four months of life with positive vocalisations increasing strongly and negative vocalisations starting high and then decreasing. The communicative nature of early infant vocalisations was underlined by the fact that different types varied across dyadic states in a meaningful way. In particular, positive vocalisations occurred most frequently during eye contact, while negative vocalisations occurred less frequently then. Parents responded with highly

diversified patterns of reactions to different infant vocalisations. Positive vocalisations were apparently interpreted as conversational behaviour and were responded to with vocal behaviours. The other types of vocalisations were interpreted as expressions of behavioural states or affect requiring parental assistance, and so the parental response aimed at changing the infant's state by tactile and vestibular behaviour. From these findings the authors concluded that parents understood the message expressed in different types of vocalisations, and reacted appropriately.

Shimura and Imaizumi (1991) focused on the role of prosodic features in the expression of emotions by two month old Japanese infants during interactive episodes with their mothers. Their study suggests that even at two months emotional contrasts are expressed on the dimensions Pleasant - Discomfort, Calm - Surprise and Singing - Speaking through modifications of certain acoustical features such as fundamental frequency range, minimum and maximum fundamental frequency values, fundamental frequency contour shape, vocalisation length and number of segments. Moreover, it was demonstrated that the prosodic patterns formed by these modifications induce consistent interpretations and responses in the caregiver(s) about the infant's emotional state. The authors note that even if the induced responses or interpretations do not correspond to the infant's actual emotional state, these are still important because they induce in the infant the idea that its own vocalisations have social meanings.

In her study of the speech of two mothers to their infants at several points between three and eighteen months, Snow (1977) found that as the infants grow older there is a steady decline in references of the mother's utterances to the infant alone, and a steady increase in utterances containing references to topics in a shared environment. The author claims that these changes in the characteristics of the mother's speech register cannot be explained as a response to the infant's need for or attempt to elicit appropriately simplified, redundant and semantically relevant speech. Rather, she proposes that the mother's speech is conveying specific information and is related in content to the messages she receives from the infant's expressions, according to her interpretation of them. This model relies on the assumption that mother-infant interaction is conversational in nature and reciprocal, that is messages are exchanges between the partners in both directions. Thus, the changes in maternal speech result from the development of the baby's increasing ability to take turns in communication (Snow, 1977).

Joanna Ryan (1974) notes that while much of what an infant utters may be difficult to understand, infant vocalisations take place within a context of interaction

with adults who are motivated to understand. Infants are actively involved in extensive communicative exchanges with their mothers during which the mother actively picks up, interprets, comments upon, extends, repeats, and sometimes misinterprets, what the infant has said. Much of what the infant learns about the language of the community is based upon the mother interpreting the infant's intent, the infant sometimes conforming with the interpretation, sometimes not but learning en route, what interpretations its efforts evoke and how these may be modified. The mother's constant interpretation of the infant's vocalisations is, certainly, one of the important factors that keep communication between them going. Parental responsiveness to infant crying or other vocal signals is given crucial significance for the formation of secure parent-infant attachment (Frodi, 1985), for affective attunement (Field, 1985) and for the parent's intuitive didactic care (Papousek and Papousek, 1987).

Although the importance of the mother's interpretation of her infant's vocalisations is widely admitted, very few studies have systematically investigated to what extent mothers rely for their interpretation on the prosodic features of the vocalisation or other situational cues. This is important information for any assessment of the meaning inherent in infant's expressions.

The Papouseks (1989) conducted a study to investigate whether single voiced units of two-month old infants which were classified beforehand by the researchers as 'cry', 'discomfort', 'comfort' and 'joy' can be distinguished when played back in the absence of contextual and other behavioural cues, and whether the ability to decode such information is determined by age, gender, parental status, experience with infants, and language. The data demonstrate that voiced sounds in the infant's presyllabic vocal repertoire effectively transmit both discrete information pertaining to the categorical distinction between comfort and discomfort, as well as graded information pertaining to the relative intensity of affective arousal. Moreover, it was found that the recipient's ability to decode state-related information from isolated test sounds varied significantly as a function of age and experience, but not in relation to parental status and gender. The mother's state estimates were influenced neither by familiarity with infant voice, nor by the cultural origin of the test sounds, nor by their own cultural origin. The authors note further that the ability to decode state related information from infant sounds is determined by an interaction of universal perceptual predispositions with experience. Experience has a negligible impact in responses to signals of severe discomfort but is crucial in responses to immature signals of joy.

Petrovitch-Bartell and her colleagues (1982) examined mothers' ability to distinguish 'fuss' and 'cry' vocalisations of their own or other three to four month old babies on the basis of certain acoustic features or other contextual cues. The results indicate that although mothers are generally able to judge 'cry' and 'fuss' vocalisations reliably using auditory information corresponding mainly to the acoustic parameters of mean intensity, mean fundamental frequency of the second formant and mean duration, they were also influenced by contextual information.

B) Infant vocalisations (and their prosodic part) are meaningful, not only because the mother's interpret them as such (Zinober and Martlew, 1985), but also because they are by nature functional in an interpersonal space, and they are functional because they are generated by intersubjective innate motives operating at a given time.

If one accepts the view that infant vocalisations are meaningful only because the mothers interpret them as such, one attributes to mothers the dominant role in communication with their babies. However, as was described before, the infant and the mother engage from the beginning of life in mutually regulated communicative exchanges, where the mother's behaviour is not only determined by her cultural background, but is also adjusted spontaneously and often unconsciously so as to match the infant's expressions, which are formed by coordinated expressive patterns in all the modalities. As Snow (1977) puts it, infant behaviours which do not have the quality of being expressive and interpretable as signals of the infant's state of mind do not seem to function in mother-infant interaction.

Halliday (1978) claims that especially after infant vocalisations have a meaning if looked at from the angle of the infant's total model of reality and of its own place in it. This reality is constantly under construction, being added to, differentiated within and modified. The author claims that long before a child begins to speak its mother's tongue, it is engaging in Acts of Meaning. For an act to have a meaning it has to be communicative i.e., to convey a message, intentional and/or symbolic. According to the author an act becomes meaningful when both the infant and the mother are taking part in a conversation which is virtually a process of meaning exchanges and the other (who is a significant other by virtue of taking part in the conversational process) joins in and gives value to the infant's acts. An act of meaning derives from a meaning potential which is a social construct. The semantic system, in which the infant encodes its subjective reality, must be shared between the infant and the significant others if his acts of meaning are to be successful. But the infant is not yet approximating the others' semantic system; it is creating one of its own. To say that the creation of a semantic

system is a social process means, therefore, at this stage, that the others must be approximating the infant's semantic system, and this is precisely what they do. The significant others actively understand the infant, they interpret and respond to its own semantic system with their own meanings. In this way the infant has access to adult meanings in a context in which they can modify and feed into its own meaning potential.

The infant's meanings are not, in general, derived from the meanings of the adult language (Halliday, 1978). No doubt, however, the adult language does exert an influence on the infant's semantic system from a very early stage, since the infant's utterances are interpreted by the communicative partners in terms of their own semantic systems. In other words, whatever the infant means, the message which gets across is one which makes sense and is translatable into the terms of the adult language. It is in this interpretation that the infant's linguistic efforts are reinforced and in this way the meanings that the infant starts out with gradually come to be adapted to the meanings of the adult language. The functional meanings that the infant expresses gradually become more and more recognisable as speech, as they come to look more and more like the meanings that are encoded into the adult language.

Locke (1996) highlights further that point, suggesting that in order to understand the steps a child takes to spoken language, one should ask: "Why do infants vocalise ?". He claims that prelinguistic vocalisations cannot be assumed to be set in motion by the anticipation of linguistic benefits. Rather, vocalisations are supposed to serve proximate or immediate functions, which are precursive to and will lead to language. In particular, infants may vocalise to share emotional experience in the form of empathy and to blend in with the individuals with whom they share an emotional relationship. By talking an infant demonstrates that it is willing and able to become a worthwhile member of a worthwhile group.

From the previous discussion it can be concluded that prosodic patterns in prelinguistic vocalisations may convey, not a linguistic meaning, but a communicative intersubjective meaning. They are part of a coherent expressive system which is functional from the beginning of life, reflecting the infant's motivational state, but also interpretable by the mother as meaningful. Thus prelinguistic prosody may express, for either the mother or the infant, an emotional state or reaction, an intention to get things done or to regulate the other's behaviour or an indication of a focus of shared interest.

2.4 On the Origins of the Form of Prosodic Patterns in Prelinguistic Vocalisations - Imitation.

The form of prosodic features exhibited in the second half of the first year may well be the result of an imitative process. A number of studies carried out to examine this possibility have come up with diverse results.

Lieberman (1967) reported that two infants produced higher average fundamental frequencies when with their mother than when with their father. A ten-month-old boy's fundamental frequency was 390Hz when with his mother and 340Hz when with his father; correspondingly a thirteen- month-old girl's fundamental frequency was 390Hz when interacting with her mother and 290Hz when interacting with her father.

Kessen, Levine, and Wendrich (1979) examined experimentally the ability of twenty three infants between three and six months of age to imitate musical pitch. An assistant was asked to present a set of standard test notes with a pitch of D, F and A above middle C either with his voice or a pitchpipe according to the infant's preference. Musically trained judges accepted as matches those responses of the baby that were within a quarter tone of the presented note. The authors observed that infants watched the experimenter closely and they often vocalised energetically; in short, they willingly participated in the task. The results demonstrated that even before the middle of the first year of life infants are able to imitate musical pitches. The authors speculate that the onset of speech might diminish the infant's ability to match pitches either because of the relatively low importance (at least in English) of precise tonal representation or because of failure of the environment to support its continuation. However, the data may have been confounded by two initial sessions during which the infants were actively trained to match pitches and thus may not reflect spontaneous vocalisations.

The most reliable evidence may come not from experiments but rather, from naturalistic studies. In a longitudinal study of ten Japanese infants from eight to twenty six weeks Masataka (1992) found that infants appear to respond to preceding maternal speech by uttering vocalisations similar in pitch contour to those exhibited by the mother. According to the author the results of the present study suggest that as the infants acquire stable laryngeal control and gradually develop communicative competence during the first half year of life, they come to alter the quality of their vocal responses according to the preceding maternal speech.

Siegel and colleagues (1990) tried to determine if infants between nine and twelve months of age would spontaneously imitate either the average fundamental frequency or

the fundamental frequency contour of their speaking partners. They also explored imitation of temporal patterns and vocal intensity. Two experiments were run: in the first experiment six girls and four boys were recorded at home as they interacted with their fathers and mothers. In the second experiment two infants were recorded while interacting with their parents in a laboratory setting. The results do not support the conclusion that infants spontaneously imitate or are influenced by the fundamental frequency of their conversational partners. Nine to twelve month old infants did not adjust vocal pitch, duration or amplitude differentially when interacting with their mothers and fathers. Furthermore, there was no indication that the infants were imitating intonation contours.

One cannot argue either that the infant's prosodic repertoire is simply the product of imitation of adult patterns or that it is completely idiosyncratic. Rather, I would agree with Locke's (1993, 1995) assertion that the infants' vocal repertoire is the product of vocal accommodation. Vocal accommodation is the process whereby infants assimilate in their own repertoire aspects of speakers' voices, speech sounds and speaking style. This assimilative process has little to do with any efforts to learn language. Rather, it appears to be motivated by the infant's desire to incorporate in its repertoire characteristics of the voice of significant others, to whom infants are emotionally attached, as a means of keeping their companionship. In other words, vocal accommodation is described as 'affect-driven' language. Vocal accommodation is a subdivision of a larger category of socially accommodative behaviours that includes gestures in various modalities. Vocal accommodation may be responsible for the first vocal behaviours that seem to be linguistic. These vocal forms encourage the infant along a path that leads to linguistic mastery by giving the infant an immediate repertoire of word and phrase-like vocal patterns (Locke, 1993, 1995).

According to Locke (1993, 1995) vocal accommodation is generated from a dual neural specialisation. This specialisation involves the specialisation in social cognition (SSC) and the grammatical analysis module (GAM). The specialisation in social cognition allows children to achieve a working vocabulary, but this specialisation is itself insufficient for an elaborate linguistic system. On the other hand, the grammatical analysis module deals with rules and representations. Thus, the SSC supplies the analytical GAM with an enormous sample of linguistically relevant data, without which there would be nothing to analyse and no forms to reconfigure by application of computational rules. However, the SSC does more than feeding the GAM with utterance data. By itself, it is also responsible for a great deal of communicative behaviours that are taken to be linguistic. The principle intake mechanisms of the SSC,

i.e. referential looking and vocal accommodation absorb a great many utterances, and with them, a variety of high-frequency words and phrases. In other words, this specialisation is slavishly vocal and socially sensitive. Thus, the earliest utterances of infants are considered to be formulaic. Formulaic utterances are described as context-bound, associative, non-computational and non-rule governed; they are holistically perceived and stored, and irreducible to their syllabic or segmental parts. Presumably formulas reflect the infant's prosodic orientation, perhaps combined with its inability to analyse speech into segment-sized elements (Locke, 1993, 1995).

2.5 The Theoretical Framework Adopted in the Present Study.

None of the preceding theories, which regard the dialogue between the mother and the infant as the basis of language development, has attempted to explain, in the infant, what generates participation this dialogue. The theory which so far has traced most successfully the foundations of mother-infant interaction and the source of its form is that of 'Intuitive Motives for Communication' (Trevarthen, 1982). According to this theory human beings are endowed with innate motives for cultural learning (Trevarthen, 1993). The motives are described as the core of the subject's intrinsic mental organisation, or otherwise as interior processes of a subject, which anticipate and interpret consequences of behavioural action, without, however, necessarily taking into account the limitations or advantages of the actual external reality for their expression. Motives are distinct from emotions, which transmit the colour of motives to the others so they can know what the subject wants. Psychological development results from changes in motives themselves, as well as from changes in the power of perceptual and motoric mechanisms which serve motives. In the course of development the fundamental motives are not always active with the same intensity. The author allows two fundamental motives: that to master objects in the environment and that to share interests and feelings with others. Subject-related evaluation or quality is given to these motives by the emotions (Trevarthen, 1982). Trevarthen distinguishes periods of development, each of which represents the attainment of a new level of cognitive and behavioural function and a new complexity of expression of the two fundamental motives. On the basis of this theory development is described as the building up of progressively more effective motives for cooperative action with other human beings.

The first or primary manifestation of the motives for intersubjective communication is the developing behaviour of 'protoconversations' (Bateson, 1975) at two months. Protoconversations are mutually regulated interpersonal exchanges between the infant and the caregiver. Detailed analysis of protoconversations has

revealed their complex organisation at the temporal and physiognomic level. In a protoconversation episode a two month old displays coordinated behavioural patterns involving all the expressive modalities to transmit distinct affective qualities which match the mother's motivational state. The infant shows sustained eye-to-eye orientation, smiles of recognition, gestures and changes of head posture, and utters animated vocalisations. Halliday (1975) notes that the infant's utterances at this stage have no semantic content, that is they refer to no specific thing, or meaning outside the direct interpersonal exchange. However, they are conversational, as Bateson (1975) claimed, in the sense of combining interest of two persons in an exchange of signs.

The period following 'Primary Intersubjectivity' extending from about twelve to thirty weeks, is characterised by clear signs of a complex and sometimes negative motivation towards intimate personal contact, but an increasing efficiency in orienting to and manipulating objects, or in attending to interesting results that others create with objects for the baby to watch. It seems that now the dominant motive is the one for exploration of the environment and gaining knowledge of the properties of objects (Trevarthen, 1982).

Because the infant often shifts attention to an object, the partner who seeks communication with the baby experiences loss of intersubjective contact. In order to engage the baby in communication, the caregiver has to display playful and dramatic behaviours. The most appropriate form of these behaviours is baby songs, or action games which become very common play routines between three and six months.

The content of baby songs, which are highly rhythmic, is constructed of a narrative sequence of feelings and expressions that has a beginning, a development, a climax and a resolution carefully paced to be startling, or exciting and so provocative of laughter (Trevarthen, 1993). The core element of the baby song is a four line stanza lasting twelve to fifteen seconds with a basic pulse of andante, a tripping rhythm, simple pitch shifts and rhyming syllables at specific points and variations in the beat to regulate excitement in the last two lines. There is increasing evidence that these elaborate features of baby songs are shared in many languages (Papousek and Papousek, 1981 Trevarthen, 1986a, 1987, 1990; Trehub, 1993). The controlled complexity of temporal patterning in these vocal action games gives evidence for a 'grammar of expression' that the infant can share long before the production of the first words (Trevarthen, 1993). The infant's muscular growth which results in better motor coordination allows it to participate in a live and enthusiastic way in these vocal action games. The infant joins in a nursery rhyme with body movements that match the

rhythm of the song and vocalisations the prosodic pattern of which complements the melody of the nursery rhyme (Trevarthen and Malloch, 1997).

The Papouseks (1981) have observed that the baby of six months who has been participating in nursery rhyme routines may express self satisfaction or joy by 'singing' when alone. The dynamic and complex sequences of communicative exchanges built up on the basis of a nursery rhyme appear to activate universal motive templates that will later serve for transmission of symbolic meanings (Trevarthen, 1993). At this age the 'topic' of communication still remains the communication itself.

Around six months after birth infants manifest increasing self-awareness (Trevarthen, 1982; Stern, 1985). This self-consciousness springs as a response to the attention of others; in other words, it is part of self-other consciousness. The sense of 'Intersubjective Self' (Stern, 1985) is manifested in showing off behaviour when the baby is either with a non supportive partner- for example when the mother keeps a still face- or with a stranger, as well as in the baby's ability to recognise itself in the mirror. At the same time the infant is gaining a more sensitive insight into the other's state of mind. The self-referred, other-sensitive emotions become stronger and clearer, and their sharing gets more subtle and better controlled (Trevarthen, 1990). The infant's increasing self-awareness facilitates the activation of the motive for defining a cooperative group with familiar persons in which meaning is created through daily practices (Trevarthen, 1993).

A primitive manifestation of person-person-object games appears at around seven months. It is very important to note that such incorporation of object exploring behaviours into interpersonal exchanges is a peculiarly human activity (Trevarthen, 1993). The infant is interested in dramatic presentations of the identities, operations and relationships between objects by the mother as well as in her emotional evaluation of these demonstrations. Further, the infant becomes an expert at seeking these habitual forms of object display and the emotional referencing accompanying them. Although the infant might passively accept the mother's participation in its activity with an object of its own curiosity or converge on the object of the mother's interest, the infant is not yet able to voluntarily adjust to the mother's interests and intentions or to direct her attention on a topic in the environment (Trevarthen, 1990). Although very little research has been done on the communicative functions of vocalisations in this period, these are supposed to express a message about the act of communication (Halliday, 1975).

Nine months is considered to be the most crucial milestone in the development communication where is essentially built the groundwork for the creation of symbols and language acquisition. The innate motives for gaining knowledge about objects and the motives for engaging in interpersonal exchanges which have been functioning separately, now go through a radical reorganisation and are combined to facilitate cooperative task purposes (Trevarthen and Hubely, 1978; Hubley and Trevarthen, 1979), and the infant shows growing awareness of how other people act as psychological beings (Stern, 1985). These changes have been characterised by Michael Tomasello as the **'Nine-Month Miracle'** (Tomasello, 1993 p.174). In particular, infants perceive the partner as a source of new ideas concerning objects, and as an intentional agent as well as a carrier of feelings (Trevarthen and Hubley, 1978; Bretherton et al., 1981). They look where the adults are looking (joint attention), they look to see how adults are feeling about a novel person or object (social referencing), they repeat what adults are doing with a novel object (imitation learning), they follow arbitrary vocal and gestural instructions, they pretend, they take initiatives in games, and they direct the other's attention to a topic of their own interest. Both partners can develop consciously joint control in the performance of a task which involves combining objects in some invented arbitrary way (Hubley and Trevarthen, 1979). On the other hand infants do not direct any of these behaviours to inanimate objects (Trevarthen, 1993). Before nine months all the infant has is its own perceptions of its direct, primary, spontaneous interactions with others and a perception of itself from the inside. After nine months the infant can perceive the adult's view of it from the outside, along with its own view of other persons from the outside. This ability allows the infant to observe similarities and differences from an outside perspective (Tomasello, 1993). Infants are now able to control the intersubjective orientation to a shared reality and convey a clear quality of interests, intentions and emotions in relation to the interest, intentions and feelings of the other; in other words infants increasingly gain 'cooperative awareness'. In previous months toys were presented in an animated way with a 'gloss' of excitement to entertain the infant; now these feelings are projected into objects in situations of 'Shared Interest' and in this way objects can gain a coded interpersonal meaning (Trevarthen, 1993).

This period was named 'Secondary Intersubjectivity' to convey the sense that now each participant is not only interacting with the other- as was the case in the period of 'Primary Intersubjectivity' at two months- but is also aware of the other's intentional participation (Trevarthen and Hubley, 1978). This is the stage at which infants gain a

sense of a 'Subjective Self' and enter the domain of intersubjective relatedness (Stern, 1985).

The developmental changes of 'Secondary Intersubjectivity' are realised as complex communicative inter-actions between the infant and the mother where meanings are exchanged. The infant contributes to these interactions with a complex repertoire which consists of combinations of forms of vocal and gestural patterns (Halliday, 1975,1978). These patterns are synchronised and equivalent (Trevarthen, 1986b, 1993). This expressive repertoire is what Halliday (1975) called 'Protolanguage'. Most infants probably use some combination of both kinds of expression, though many show preference for one particular kind (Halliday, 1975, 1978).

CHAPTER 3

3.1 The Aims of the Present Study

The literature review has revealed theoretical and methodological gaps in the study of infant prosodic patterns in the second half of the first year. So far, relevant researches have only sought for similarities and differences between infant and adult prosodic patterns. What is missing is an ethological study which would study infant prosodic repertoire at the threshold of language in its own right, and interpret its psychological role in mother-infant communication, the essential context for language development, . Moreover, one can note that studies on the role of infant prosodic patterns in conveying messages towards the end of the first year have focused on communicative functions, ignoring the emotions that saturate and qualify every communicative exchange. Considering the above deficiencies this thesis aimed to complete:

1) A longitudinal description of infant prosodic patterns over the transitional period to Secondary Intersubjectivity, i.e. between thirty and fifty weeks. It is at this age that human infants can begin to engage in cultural learning since in cultural learning learners do not just direct their attention to an individual and its behaviour, they actually attempt to see the world the way the other individual sees it, from inside the other's perspective (Tomasello, 1993).

2) An analysis which would not search for any formal or functional continuities between prelinguistic prosodic patterns and the ones appearing in adult language, but rather one that would go a step backwards and study the former in their own right, investigating their role in relation to the other modalities in mother-infant communication, which constitutes the groundwork of language development.

3) A methodological approach which would specify the meaning in infant vocalisations as adequately as possible taking into account all the parameters which identify it.

4) An investigation of the forms and variations of infant vocalisations that convey not only communicative functions but also emotions.

In sum, the purpose of the present study is to investigate the role of contrastive prosodic patterns in infant vocalisations in expressing different emotional or mental states, or serving different communicative functions in relation to the other expressive modalities (i.e. gaze, gestures, postures).

DATA COLLECTION

3.2 The Problem of Attributing Meaning to Infant Vocalisations

3.2.1 Introduction

Depending on the theoretical framework adopted, studies of infant vocalisations have dealt with the issue of attribution of meaning to these vocalisations in a variety of ways.

Followers of the linguistic approach have attributed meanings to infant vocalisations using categories devised for the description of meaning in the intonation patterns of adult language However, as it was discussed in Chapter 2 if one is attempting to find out how linguistic expression develops, the emerging infant meaning system cannot be described with categories defined for the already developed language system.

Some researchers have identified the meaning of infant vocalisations in terms of a functional coding of accompanying gestures (Dore et al, 1976; Harding and Golinkoff, 1979; Marcos, 1987; Flax et al, 1991). For example, a vocalisation is classified as 'Giving' if uttered while the infant is holding out an object to the mother, and a 'Request' is identified as a vocalisation combined with the hand stretching after a desired object has been taken away by the mother. However, this method is ambiguous because, although infant gestural and vocal repertoire are combined in a coherent expressive system, it might be the case that at a given time these two communicative channels convey different but yet not contradictory messages. To continue with one of the previous examples, hand stretching towards the desired object after this has been taken away may signal 'Request' for the object, while the accompanying vocalisation may express frustration for not having it. A given psychological state of action does not consist exclusively either of emotional or of referential content; rather, these two aspects of the motive for action coexist, and in each case the relation between the emotion and the referring is reflected in the expression. If one aspect of the psychological state is dominant, then it is more probable that this one will be expressed simultaneously by both gestural and vocal channels. If the emotional and the referential aspect are of equal strength, then they may be expressed in different modalities. Furthermore, when studying the development of an expressive system one cannot assume that message categories can be determined beforehand; and moreover, one cannot know at which

stage of development particular gestural or vocal forms will be combined to convey the same message.

A third way of inferring the meaning of a particular infant utterance would be to analyse the mother's immediate vocal or non-vocal response to it. Although this method would undoubtedly provide useful information about the network of mother-infant communication, it is not reliable for the following reasons:
(a) Mothers do not always actively respond to their infants' vocalisations, (b) Mothers may respond with the same behaviour to infant vocalisations conveying different messages e.g. they may give the infant an object after a vocalisation expressing 'Frustration' or 'Request', furthermore, (c) this method does not allow for an interpretation of vocalisations occurring when the infant is playing alone.

None of the above three ways of defining the meaning an infant intends when vocalising was suitable for this thesis, the principal aims of which were to discover how vocal expression is used by an infant to establish intersubjective communication with the mother, or to express psychological state, and to relate different expressions to the context of behaviour. It was therefore decided to adopt a method of analysis that related three independent ways of describing the characteristics of the infants' utterances without making theoretical presuppositions about their functions.

3.2.2 The Method of Attributing Meaning to Infant Vocalisations followed in the Present Study.

Examining the meaning of animal vocalisations, Marler and his colleagues (1992) noted that, even though we know that most humans ultimately will develop language, infants, like non-human animals, cannot be directly interrogated about what they think, mean, or intend. Thus, the developmental linguists and the ethologists face common theoretical and methodological obstacles in establishing the meaning of a vocal utterance (Marler et al, 1992). The authors conclude that the problem of meaning attribution should be approached ethologically, i.e. the researcher needs to empirically establish correlations between vocalisations and emotions, mental states or communicative functions, without being guided by any predetermined theoretical assumptions. The principle point of ethology is that any theory about behaviour, especially with respect to its evolutionary origins, should rely on exhaustive lists of precise, operationally defined behaviours in naturalistic settings (Lorentz, 1966; Hess, 1970).

I would agree with this method, and would add that a method for attribution of meaning to infant vocalisations must ultimately consist of an insightful investigation of the situation in which these meanings are generated that is, mother-infant communication. An essential part of this communicative process is the interpretation of infant vocalisations by the mother. In Chapter 2 (Section 3) it was argued that one of the factors which render infant vocalisations meaningful is that mothers spontaneously, and often unconsciously, interpret them as having meaning. Although a mother may not at all times overtly respond to her infant's vocalisations, she always intuitively interprets these vocalisations, even the non-communicative ones, and this facilitates the communication of the dyad.

In the interpretation of any infant utterance by the mother the following different features are also involved in addition to prosody:

- The non-vocal expressive accompaniments of the utterance in the infant's behaviour.

- The circumstances of the utterance, i.e. the presence or absence of particular objects, or any other relevant change in the physical environment related to a vocalisation,

- Presuppositions made about the communication; that is, any information assumed to be shared by the infant and the mother.

It is obvious that only the mother has access to all this information necessary for the attribution of a message to her infant's vocalisations. Thus, it was thought that the first step in a method for attributing meaning to infant vocalisations should be to ask the mothers themselves what, if anything, they feel their infant is expressing in each instance.

3.3 The Pilot Study

A Pilot Study was undertaken to help in the design of the Main Study, testing theoretical and methodological preparations for the naturalistic study of infant vocalisations. These prior considerations can be summarised as follows:

(a) Prosody in infant vocalisations is part of a coherent system including all the expressive modalities.

(b) The infant vocal repertoire, rather than being simply a precursor to the adult language, is an expressive system in its own right, serving communication with what the

mother feels, experiences and intends, as well as conveying the infant's purposes, intends and feelings.

(c) The mothers tend to intuitively read meaning in most of their infant's vocalisations, even the non-communicative, self-directed ones.

The Pilot study served the following purposes:

- to gain experience using the recording equipment;
- to decide on the age of the participants in the Main Study;
- to decide about the most appropriate materials and situations, and the duration of filming session that would give adequate information for data analysis;
- to form the questions for the mothers, seeking information about what they feel their infant means, if anything, in each and every vocalisation;
- to ensure that the mothers would respond to the interview in a consistent and similar manner, so as to be able to compare their answers;
- to obtain a sample of audio and video tapes on which to practice methods of analysis.

In the cros-sectional Pilot Study 16 infants and their mothers, recruited from the departmental list of volunteers, participated. All infants were first borns, except one boy who had two older siblings. They were born to middle class, two-parent, English speaking families; only one boy was born to a single-parent family. Four cohorts of mother-infant dyads were observed over a 30 week period, from 26 to 56 weeks. Each cohort involved 4 infants, 2 boys and 2 girls, with the exception of the eldest cohort which included 3 boys and 1 girl. Dyads were filmed at their homes every fortnight for two months. Ten out of the sixteen mothers also accepted to participate in Pilot interviews in the Department on the meaning of their infant's vocalisations.

How the observations in the Pilot Study helped in the design of the Main Study is described in relevant parts of the following Section.

3.4 The Design of the Main Study

3.4.1 Subjects

Participants in the Main Study were two first-born infants, a boy and a girl, and their mothers both from middle class, two-parent, English speaking families. The subjects were recruited from the departmental list of volunteers. First-borns were selected for two reasons: 1) It was felt desirable to include mothers who were likely to be naive about any changes that might occur in the infants' behaviour during the period of study, and 2) It was decided to avoid any questions concerning the effects a sibling might have on the infants' language development.

The infants were studied longitudinally over the period from 30 to 50 weeks. This age range was chosen because it was felt necessary to study the role of infant prosody - a feature which so far has been studied mainly from a linguistic viewpoint - from a non-linguistic perspective over a pre-verbal period when psychological transformations occur that are crucial for the subsequent development of language; namely, the coordination of the infant's own feelings and interests with those of the mother. A detailed description of these transformations is presented in Chapter 2. The study did not continue beyond 50 weeks because, as demonstrated in the Pilot Study and in accordance with previous relevant studies, most children start producing recognisable words at this age (Vihman et al, 1985), and the purpose was to focus on prelinguistic processes.

3.4.2 Recording Equipment

For the video recordings, a Panasonic HiFi camera was used, mounted on a tripod. In order to provide signals of good enough quality for subsequent analysis of prosody, additional audio recordings were made simultaneously using a SONY Digital Audio Recorder, with an external one-channel microphone. Lapel microphones attached to the mother's and the infant's clothing were not used because it was found in the Pilot Study that these can distract an infant's attention or obstruct his or her movements. The microphone was mounted on a tripod placed next to the camera. The mother was asked to try to be sitting on the floor opposite to her infant, in position such that camera and microphone would be aimed to the space between them.

3.4.3 Video Recordings

3.4.3a Location

The recording sessions took place in the infants' homes every fortnight for five months in the room where the baby and the mother were used to play, and during a time when the mother felt her baby would be awake and alert.

In many studies on infant vocal development recordings have been run in laboratory environments, presuming that speech samples obtained will be representative of the infants' optimal vocal performance (Delack and Fowlow, 1978; de Boysson-Bardies, 1982; Furrow, 1984; Marcos, 1987; D'Odorico et al, 1991). However, studying the effects of different recording environments, specifically laboratory vs. home settings, on speech samples of typically developing infants, Lewedag, Oller and Lynch (1994) found that infants vocalise more than twice as much in the home environment. In addition, it was shown that infants produce a larger proportion of both canonical babbling (an 'age immature' behaviour) and vowel sounds (an 'age mature' behaviour) in the home than in the laboratory. The results of this study also suggest that vocalisation samples collected in the laboratory may tend to under-represent the level of vocal maturity of the infant shortly after the onset of canonical babbling (i.e. in the period from 4.5 to 7 months) (Lewedag et al, 1994).

The home environment was considered to be more appropriate than the unfamiliar laboratory, for observations of how the mother and the infant express themselves spontaneously. Moreover, in a naturalistic environment, unlike a sound-treated laboratory room, it is not always possible to avoid background noise from sources irrelevant to the session. Nevertheless, nearly all audio recordings obtained in this study were of very good quality.

3.4.3b The Video Recording Situations

In order to meet the purpose of the present study, it was important to collect from each and every recording session a wide range of behaviours exhibited in various contexts of action and interaction, and there would necessarily include a representative sample of each infant's vocal repertoire.

It was clear that the recording set up should allow the infant and the mother to express themselves as they might do in daily life, whilst encouraging a large variety of spontaneous performances. The Pilot Study showed that unstructured sessions do not

ensure the occurrence of a variety of play activities and situations. It was often the case that the infant and the mother were engaged in one and the same activity for a long period of time. Thus, it was decided that a structured naturalistic procedure would be the most appropriate one for collecting a substantial corpus of vocalisations from both infants. The method is characterised as 'naturalistic' since the recordings take place in the infants' homes and in familiar situations of everyday life. It is characterised as 'structured' because the introduction of certain situations by the researcher did not leave the dyad complete freedom in the development of play activities. The structured naturalistic recording method also ensured that both infants would experience similar situations during the sampling period, thus, favouring inter-infant comparisons.

The above considerations led to the following plan for a sequence of activities which were offered to the mother in the recording situations.

1. ***Free -play interaction:*** In this situation the mother and the infant were free to interact in any way they themselves would choose. Usually the mother or the baby would introduce a toy, and both became engaged in shared action. They explored communicative routines that had been developed in daily life, routines which eventually lead to the acquisition of language (Bruner, 1975; Trevarthen, 1990).

A wide range of communicative cooperative and isolated behaviours was favoured by imposition of the following pre-planned situations:

2. ***Mother-Infant play without using toys:*** This situation was included to encourage playful person-person communication, giving the dyads the opportunity to play musical or other body games that might exist in their spontaneous repertoire and of a kind that has been shown to elicit various types of vocal behaviour (Papousek and Papousek, 1981). The mothers were instructed to engage their infant in any play activity but not to use any toys or other objects.

3. ***Infant Playing Alone:*** Solitary activity by the infant has been demonstrated to be accompanied by rich vocal activity. The mother was asked to move into another room and stay outside her infant's range of vision until she was called back.

4. ***Mother-Infant playing with researcher's toys:*** This situation gives both infants the opportunity to play with the same set of toys. Three toys were supplied: a music box which the infant could not wind up, a wooden multicoloured puppet figure whose legs and arms could be moved by pulling a string to create an interesting effect for the baby to watch, and a brightly coloured leather ball which could stimulate a 'give and take'

game between the mother and the infant. These or similar kinds of toys have been used in previous studies to elicit particular types of vocalisations (Marcos, 1987) or sets of intentional behaviours that include vocalisations (Carpenter et al, 1983; Hubley, 1983). Neither of the babies had seen these toys before, and they were not presented with similar objects during the period of the study. The toys were kept in a box and the mother was instructed to choose one at a time at different points during the session, in order to avoid the distraction of a plethora of new toys.

5. ***'Car' Task :*** This task has been used in previous research to study the development of cooperative action in the second half of the first year, to provoke directive behaviour from mothers and to permit measurement of changes in the infants' responses, including vocalisations, with age. The findings have shown that the infant's abilities for cooperative action develop dramatically after the age of 40 weeks, constituting one of the major psychological achievements of the first year (Trevarthen and Hubley, 1978; Hubley and Trevarthen, 1979; Hubley, 1983). Given this fact, the 'Car-Task' was felt to be an appropriate paradigm for studying vocalisations uttered during cooperative activity. A wooden car that could hold four different coloured wooden dolls and a toy dog was presented to the mother who was asked to teach her baby to put the little dolls and the toy dog into the car.

6. ***'Modified Still-Face' Task:*** This paradigm deviates from the naturalistic character of the other conditions, but it was chosen because it creates an age appropriate mildly stressful situation which has been demonstrated to elicit a wide range of affective behaviours from infants, vocalisations included (Murray, 1980; Murray and Trevarthen, 1985; Weinbergh and Tronick, 1994). At some point during a play activity the mother was instructed to pretend that she was reading something and ignoring her baby.

Each session lasted approximately thirty minutes, starting with free play interaction between the infant and the mother. At some point the mother was instructed to change to another situation specified by the researcher. The precise manner of the transition from one situation to another was controlled by the mother. Each situation lasted approximately three to five minutes. If the infant was not cooperative or persistently showed resistance to the new situation (e.g. crying or following the mother as soon as she left the room) this condition was terminated and an attempt was made to introduce it again later in the session. Situations (2), (3), (4), (5) and (6) were applied only once in each recording session whilst situation (1), Free Play, was applied more than once, each time for approximately three minutes. The order of the observational conditions varied across infants and sessions, so as to counter possible fatigue effects.

Every effort was made to include all the above situations in every recording session; however, this was not always possible. In particular, it was observed that as they became older both infants spent more time in coordinated joint play, less time in person-person activity and less time unengaged with a person, or unoccupied with an object or an activity, and they increasingly exhibited great resistance every time the researcher asked the mother to leave the room or to play with her baby without using any toys. These observations are in agreement with findings reported by Bakeman and Adamson (1984), who studied longitudinally how six-to-eighteen-month-olds coordinate their attention to people and objects in mother-infant and peer-infant interaction. As a result of these developments, it was either impossible to apply the situations 'Mother-Infant playing without toys' and 'Infant Alone' in the last sessions, or they could be applied for less than three minutes. Nevertheless, this was not considered to be a problem, since it simply reflects a normal developmental trend. Moreover, it is worth noting again that the purpose of this study is not to find if babies utter particular types of vocalisations in certain situations. Rather, the different situations are applied in order to trigger any augmentation of spontaneous vocal performance.

The researcher was present in the room during the recording session and both infants occasionally communicated vocally with her or about her. These vocalisations were included in the sample since they consist of a spontaneous reaction to a contextual event.

3.4.4 The Mothers' Interviews

The interviews took place in the Department within a few days after each filming session, and before the next one, when convenient for the mothers. For their transportation a taxi was arranged and paid by the researcher.

First the use of the remote control of the video player was demonstrated, and then the following written guidelines were given to the mother:

Part I

"In a few minutes you are going to watch yourself playing with ( ), or ( ) playing alone. When you feel that ( ) says something, please press the button 'stop' and tell the researcher as soon as possible what ( ) is expressing or saying. In case you feel that ( ) expresses more than one message, imitates you, or just plays with his/her sounds, please report it. Please use one word or a short phrase if at all possible".

Part II

"Then you have as much time as you need to describe what is happening, what ( ) is trying to do, and who ( ) is talking to. When you have finished answering all the questions, press the 'play' button to view the rest of the tape. Please speak clearly and the microphone will record your voice".

The first part of the interview was designed to reduce the role of rational decisions and to increase the probability of intuitive responses, whilst the second part aimed to highlight the answers given in the first part. Each interview session lasted approximately one hour. The mothers responded to the interviews enthusiastically.

An alternative to the form of interview described above would be to play back just the audio tape and ask the mother's to describe the message they feel is conveyed in each of their infant's vocalisations. This method has been followed in studies on the acoustic determinants of responsiveness to infant vocal expressions (M. Papousek, 1989; Shimura and Imaizumi, 1991). However, this reliance on a single modality is not appropriate for studying the role of infant prosody in relation to the other expressions used to convey meanings. As it was noted earlier in this Chapter (Section 4.1.2), in natural communication the mother relies not only on prosody, but also takes account of many other characteristics of the situation when interpreting her infant's vocalisations.

Thus, interviews aiming to investigate the mothers' interpretations in an ethological way need to simulate the natural conditions of communication.

The independent analysis of prosodic characteristics in infant vocalisations, the related infant non vocal behaviour, and the mothers' interviews will together reveal to what extent the infant distinguishes, and consequently the mother interprets, individual messages on the basis of prosody or other expressive behaviours.

CHAPTER 4

DATA ANALYSIS

In the present thesis three kinds of analysis were carried out: (a) functional analysis of the dyad's behaviour related to each infant's vocalisation, (b) analysis of the mothers' interpretations of their infants' vocalisations, and (c) analysis of the prosodic features in each and every infant vocalisation.

4.1 Functional Analysis from Videos of the Situation of Infant Vocalisations

The purpose of the present study is to examine the role of infant prosody in conveying meanings and to assess the role of other non-vocal expressive modalities; that is, gaze, gestures, postures and facial expressions. As was described in Chapter 4, the meaning of every infant vocalisation in this sample was in the first place identified by the mothers themselves. In order to assess the contribution of prosody to expressed meanings, ultimately one needs to investigate the extent to which the mothers relied for their answers on prosody or on other characteristics of the situation in which a vocalisation occurred. Therefore, each infant vocalisation was independently interpreted by the researcher attributing a function or an emotion to the accompanying expressive non-vocal behaviour, whilst noting other relevant features of the situation, such as the timing of the vocalisation in relation to a change of an 'active' object. This independent attribution of the message of each utterance, made by microanalysis of the videos, was then compared with the mother's interpretation.

In cases where, according to the analysis of prosody, vocalisations displayed the same prosodic pattern, but were interpreted as different in meanings by the mother, and where the coding of the non-vocal behaviour from the video record is in agreement with the mother's interpretation that the vocalisations have different meanings, then one can conclude that the different messages are mainly conveyed through the non-vocal behaviour, or that they are inferred by the mother from other situational cues, rather than prosody. On the contrary, if vocalisations attributed different messages are accompanied by non-vocal behaviours serving the same function, or they share the same situational characteristics, but display distinct prosodic patterns, it can be concluded that in this case prosody is the prime carrier of meaning.

The term 'function' is here being used in the way proposed by Halliday (1975) in his analysis of the 'Protolanguage' of an infant aged nine to eighteen months. Halliday

suggested that, in its developmental origins, the linguistic 'function' of an utterance was synonymous with its 'use'. The function of a non-vocal behaviour observed in the present analysis was therefore taken to be the use it served in the situation of its occurrence. A behaviour was attributed a function if its purpose or use was clear.

4.1.1 Defining an 'Episode' in which an Infant Vocalisation occurs

In their study of selective production of prosodic patterns in different communication contexts D'Odorico and Franco (1991), identified the communicative context by glances or gestures occurring either during the vocalisation, or in the first two seconds immediately following it. However, pilot analyses made in the present study showed that a vocalisation cannot be attributed a message on the basis of a behaviour preceding or following it, because at this age the changes in the infant's focus of attention, communicative intention and emotional state can be so rapid that the non-vocal behaviour surrounding a vocalisation may convey quite different message(s) from the target vocalisation. It was, therefore, considered more reliable to attribute a function or an emotion to a particular vocalisation only by reference to any expressive non-vocal behaviour that is coincident with this vocalisation.

In any attempt to describe the function of human behaviour one must take into account the fact that behavioural units do not occur individually, as separate events, but rather are part of sequences, which form psychological episodes. Thus, the function of an expressive non-vocal behaviour can be described reliably only in relation to the function of the rest expressive non-vocal behaviours in the episode in which it occurs.

Episodes were identified as temporal events. A communicative episode could not be defined functionally, since it was itself to be the framework for functional coding of the non-vocal behaviour occurring within it. In the age studied that is, from 30 to 50 weeks, vocalisations are presumed to be related to the present situation, and they do not refer to a chronologically distant event. It has been demonstrated that attention cycles in infant-caregiver interaction occur within three to six seconds (Kaye and Fogel, 1980). Also, the three to six second temporal period has been described as possibly fundamental to human motoric and perceptual function (Fraisse, 1982). This length of time is also in accordance with the normal duration of a linguistic phrase, a line of poetry, or a musical phrase (Pöppel, 1994). A vocalisation may occur at the beginning in the middle or at the end of a communicative episode, or to mark the transition from a communicative exchange to another. Thus, it was felt that the function of a vocalisation in relation to the ongoing behaviour can be reliably studied in an interval

defined six seconds before the beginning and six seconds after the end of the vocalisation. The coding of the visible behaviour was made without sound and the onset and offset time of each and every behaviour was also noted, so as to record the temporal sequence of the communication. In cases the non-vocal behaviours revealed both a function and an emotion, e.g. when the infant was smiling while offering an object to the mother, both messages were noted.

4.1.2 'Expression' in Non-Vocal Behaviours

An issue that rises at this point concerns which non-vocal behaviours can be considered as 'expressive'. As Wiener and his co-authors have noted (1972), the term 'expressive non-vocal behaviour' can get a twofold explanation: it can either include a person's actions which an observer interprets as conveying thoughts or feelings, or it can mean that a behaviour is a manifestation of the internal state of the actor. The first explanation is mainly concerned with an observer who makes inferences about the concurrent cognitive or affective experiences of the actor. On the other hand, the second explanation shifts the focus to the subject who intentionally or unintentionally conveys his or her own present experiences through the non-vocal behaviour. If one accepts the first explanation, one is obliged to include in the term 'expression' only behaviours which are observed by somebody to be expressive, and to exclude others which may express the person's present interest or emotional state but that occur while the person is alone. In this study 'expressive non-vocal behaviours' were considered to be all the non-vocal behaviours that were apparently not caused by physical factors and were not reflexes. Expressive non-vocal behaviours were defined as 'communicative' when they were evidently directed to the interests or actions of another person.

The form of an expressive action is important, because different messages may be conveyed by changing the way an expression is performed. However, for communicative non-vocal behaviours the form alone is not sufficient to convey meaning. Actions of similar form may gain different communicative values or perform different functions, according to the partner's preceding action or response. For example, when 'holding out an object' to the mother in response to a request, the infant may be agreeing to share interest in the object with her. On the other hand, when a 'holding out' action is performed spontaneously it is possible that the infant's intention is to attract the mother's attention to the object. The first 'holding out' action can be characterised as 'responsive', while the latter is an 'initiative'. Thus, an infant non-vocal behaviour can reliably be attributed a meaning only after an insightful examination of its relation to the functions of the expressive non-vocal behaviours of the

communicative partner. To fulfil this requirement, not only the infant's non-vocal expressive behaviours, but all the mother's expressive non-vocal behaviours exhibited in a communicative episode were also analysed frame-by-frame and codified functionally.

4.1.3 The Code of Functions Developed for Analysis of the Video Records

At the start of this study little was known about the full range of functions each infant could convey through its non-vocal behaviour, since it is concerned with a developing expressive system. So it was felt necessary to begin the functional classification of the infant non-vocal behaviour by defining refined functional descriptive categories from the observation of the behaviour itself, rather than by using functional categories based on previous reports and defined *a priori*, such as 'pointing', 'showing' and 'giving' already known to be used by infants in the second half of the first year. Taking into account the above considerations, a system of 71 mutually exclusive functional descriptive categories were defined (Appendix I). These categories were then grouped into three classes: 'intersubjective behaviours', 'solitary behaviours' or 'emotions':

I) Intersubjective Behaviours

I) Person - Person (Interpersonal): The functional classes included in this group comprise categories of behaviour addressed to the mother or the stranger without involving any object.

a) Response to Invitation to Play
b) Invitation to Mother to Play
c) Attention Seeking
d) Invitation to Stranger
e) Request for Action on Self
f) Conflict
g) Showing Off

II) Person - Person - Object: This group includes functional classes comprises categories of behaviour involving the mother and an object.

h) Converging Interest
i) Joint Interest
j) Directive

The functional class 'Converging Interest' includes behaviours which manifest the infant's passive acceptance of the mother's participation in activity with an object of the infant's own curiosity, or when the infant's interest simply converges on the object of the mother's interest. On the other hand, the classes 'Joint Interest' and 'Directive' comprise behaviours, which become common after nine months, that reflect the infant's increasing ability to voluntary adjust to the mother's interests or intentions on objects, or to direct her attention towards a topic in the environment.

k) Request for Object
l) Conflict on Object
m) Split Attention

2) *Solitary Behaviours*

Non - Shared Interest in Object.

3) *Emotions*

a) Shared Positive
b) Shared Negative
c) Non-Shared Positive
d) Non-Shared Negative

The infant was thought to share an emotion, Positive or Negative, if the facial expressions conveying this emotion was accompanied by other communicative behaviours.

At a given time, more than one form of expression may contribute to the transmission of a particular message. However, it can be the case that not all of them convey functions of the same psychological complexity. Rather a particular modality may convey functions which reflect greater psychological complexity than another. For example, in a situation where the baby looks at the mother's hands, while spontaneously holding out an object to her, and does not resist when she takes the object (*Offers*), the gaze behaviour itself does not show anything more than the fact that the baby Seeks Contact with the mother. It is the gesture that reveals that the baby's motivation at this moment is not simply to Seek Contact with the mother but to Direct her Attention to the object offered. It would seem reasonable, in such a case, to classify the infant's motivation at a given moment by the behaviour(s) which reflects the more psychologically complex function.

143 minutes of video recording (64 for the boy and 79 for the girl) were codified in the way described above.

4.2 Analysis of the Mothers' Interviews

The mothers' answers in the interviews about what they feel their infant was expressing in each vocalisation were brief, precise and described similar messages showing also the same developmental trends. Some of the most common answers were the following:

Boy

7 months

1) "Interest. He was looking down at the toy. I don't know what he was trying to do with that, but he was just telling me that he was interested in playing with that".
2) "He is trying to get my attention. I was reading a book. He is talking to me".

8 months

3) "Delighted. I'm showing him the 'dancing man'".

9 months

4) "Irritation. The music stopped; he is still exploring the possibilities of the toy when the music is playing, so he doesn't like when the music stops because he hasn't figured out what to do with it. He is talking to me".

10 months

5) "Recognition. One of his friends has the same ball with different colours".
6) "He noticed the cup and the ball. Request to me to start a game. He knows that I'm going to hold the cup and he is going to put the ball in it".

11 months

7) "He is talking to me. The music stopped. He is asking for the music to go on".

Girl

7 months

8) "We are playing pick-a-boo. I opened my hands and she uttered a squeal of recognition and pleasure that I was there. Talking to me".

8 months

9) "Enjoyment. She is playing around and shaking the toy. She is talking to me saying that she is having fun".

10) "Displeasure. She is talking to me. I took her rattling toy away from her and this is a whimper".

9 months

11) "Greeting for 'Agi', her teddy bear".

12) "Conversational noise. She is talking to me. She is telling me a story possibly about the teddy".

10 months

13) "She is giving me the toy. She is showing it to me. She has been playing with the toy and I've been playing with her. She is picking it up and it is as if she is saying "Look". She is pushing it in my hand and she is taking it away as if she is drawing my attention to it".

11 months

14) "She is responding to what I'm saying".

In most cases the classification of the mothers' answers retained the terms the mothers themselves employed to characterise the messages of their infants' vocalisations. These terms clearly described mental states (e.g. interest, recognition), communicative functions (e.g. requests, directives), or emotions (e.g. pleasure, frustration). In cases of periphrastic attributions, these were still explicit enough to be classified in one of the above categories (see example 13). Vocalisations described by the mothers as 'Conversations' or 'Stories' were classified as 'Narratives'. Whenever the mothers reported that a vocalisation was addressed to her, the message this vocalisation conveyed was also classified as 'Shared'.

Table 4.1 shows the correspondence between the terms used for the classification of the mothers' answers and the categories devised for the video analysis.

Table 4.1: Correspondence between the terms used for the classification of the mothers' answers and the categories devised for the video analysis.

	Categories **used** in the Video Analysis	Categories **not used** in the Video Analysis
Categories **used** for Interviews	Attention Seeking Invitation to Stranger Request for Action on Self Request for Object Directive Shared Positive Emotion Shared Negative Emotion Non-Shared Interest	Shared Narrative Shared Comment Response Acknowledgement of Presence Request for Action on Object Shared Recognition of Object Non-Shared Narrative Practising
Categories **not used** for Interviews	Response to Invitation to Play Invitation to Mother to Play Showing Off Conflict Conflict on Object Split Attention	

The messages 'Shared Narrative', 'Non-Shared Narrative', 'Shared Comment', 'Response', 'Shared Recognition of Object' and 'Request for Action on Object', which were reported by the mothers, could not have been identified by analysis of the video records. Regarding the message 'Request for Action on Object', neither of the infants accompanied the vocalisations interpreted as such with a non-vocal behaviour conveying the same message, e.g. giving an object to the mother after trying unsuccessfully to perform an action with it. In his study on the prerequisites of language, Bruner (1982, 1983) observed non-vocal behaviours conveying 'Request for Action on Object' after 13 months. The author notes that for requesting assistance in carrying out a goal-directed action, the infant must combine his/her knowledge for means-end relations in the real world with the communicative procedures for gaining help in executing them. Knowing how to ask for help implies knowing at least some of the arguments of action involved in the task, at least in a conceptual sense. It is not surprising then that early supportive requests depend heavily for their success upon adult interpretation (Bruner, 1983). The cues on which the mothers relied for the interpretation of their infants' vocalisations as 'Requests for Action on Object' will be analytically presented in the Results' Chapter.

The messages that involved the infant and an object, or the infant, the mother and an object, were characterised as *'Referential'*, while the messages that involved the mother and the infant only were characterised as *'Interpersonal'*. Another distinction is made between *'Assertive'* messages which express the infant's intention to involve another person in the play, or to seek for information from the environment, and *'Receptive'* messages which express acceptance of a given situation or taking in stimuli from the environment.

The mothers of both infants used a variety of terms in the description of positive and negative emotions including 'Happiness', 'Enjoyment', 'Delight', or 'Pleasure', and 'Unhappiness', 'Irritation' or 'Frustration' respectively. However, Plutchik (1980) has shown that the terms of each emotional category referred to above, are not semantically differentiated. So, the vocalisations characterised by different terms describing each of the two kinds of emotion were classified in one category, i.e. as 'positive' or 'negative'

The messages reported by both mothers can be classified as follows:

A. Assertive
1) Interpersonal: 'Attention Seeking', 'Invitation to Stranger'.
2) Referential
i) Other-Directed: 'Shared Interest', 'Directive'.
ii) Self-Directed: 'Non-Shared Interest'.
3) 'Request'[1]

B. Receptive
1) Interpersonal: 'Acknowledgement of Presence'
2) Referential
i) Other-Directed: Shared Recognition of Object
ii) Self-Directed: 'Practising'
3) 'Shared Narrative', 'Shared Comment', 'Response'
4) 'Non-Shared Narrative'

C. Emotional
1) 'Shared Positive Emotion'
2) 'Shared Negative Emotion'

[1] In the analysis of prosody the different categories of 'Request' were collapsed in one category.

In few of the mothers' answers there was an element of exaggeration, which, however, was not located in the attribution of message as such, but in any further explanation or description. This element of exaggeration amplified the main message. For example: Boy 11 months: *"Recognition of the toy. It is a sort of 'Do you remember that mummy?"*. Girl 8 months: *"Give me the ball. Don't be cruel!"*.

Careful examination of the mothers' answers in the Main and the Pilot study showed that the mothers did not attribute messages that their infant would not be likely to convey at that age. One can conclude that the mothers of this sample interpreted their infant's vocalisations in a common and systematic manner which, moreover, is in accordance with the infant's developmental stage. It seems that, notwithstanding individual differences, communication in different mother-infant dyads is organised on a common groundwork in a way that facilitates infant's expression.

4.3 Analysis of Prosody

4.3.1 Introduction

In Chapter 1 prosody was defined as 'any perceivable modulation in pitch or loudness over time that conveys meaning'.

Infant vocalisations, like speech utterances, are complex sounds, i.e. they consist of combinations of many simple sinusoid waveforms with different frequencies; the lowest of these frequencies is called fundamental frequency (F_O). The fundamental frequency represents the rate of vibration of the vocal folds during the voicing of an utterance, and it is the acoustic correlate of the perceptual feature of pitch. Loudness is the perceptual feature relating to the physical concept of intensity. Intensity is proportional to the square of the amplitude of oscillations of air molecules in sound-waves passing through the atmosphere.

4.3.2 Methods of Analysis of Prosody in Infant Vocalisations

Going through the relevant literature, one notices a diversity in the ways that prosodic features in infant vocalisations have been analysed. These methods reflect the theoretical views adopted as regards the nature of the phenomena studied.

a) The Acoustic Method

The 'acoustic' or 'concrete' approach to prosodic analysis relies on the definition of prosody in purely physical terms, as those phenomena that involve the dimensions of time, frequency and amplitude (Ladd and Cuttler, 1982) . In each of these domains a number of features have been studied using computer programs for sound analysis. In addition, combined parameters, consisting of features from two or more domains can be constructed (e.g. the 'formants' which are, the energy peaks at particular frequencies).

Some researchers have extracted fundamental frequency measurements of infant vocalisations visually from the spectrogram (D' Odorico and Franco, 1991, Robb and Saxman, 1985). The spectrogram provides an impartial record of the general physical condition of a vocalisation, i.e., the energy of various frequency components in the vocalisation as well as the amount of noise embedded in it. This latter information is very useful in preparation for making a pitch extraction. However, one cannot obtain precise fundamental frequency values from a spectrogram. Reading of a spectrogram record is an art rather than a reliable objective method of measurement (Scherer, 1982).

Some researchers suggest that the appropriate way of obtaining fundamental frequency values is to generate period-by-period fundamental frequency plots over a specified time window (Sheppard and Lane, 1968; Laufer and Horii, 1977). Objective measurements of fundamental frequency provide useful information on the acoustic level about the characteristics of the speech wave form emanating from the mouth. However, these measurements do not represent in any way what can be perceived by the hearer, which is a crucial information for a study that focuses on communication processes. This is due to the fact that, although fundamental frequency is the physical correlate of the perceptual variable 'pitch', the fundamental frequency of complex tones does not, in general, vary linearly with pitch. For sensation levels from 5 to 60 dB, pitch discrimination is poorer for low tones, and best above 1000 Hz (Seashore, 1967).

b) Subjective Assessments

Many well-known studies which investigated the role of infant prosody in conveying meaning in naturalistic mother-infant interaction have focused exclusively on the final pitch contour movement, this being assessed impressionistically as 'rising' or 'falling' (Darwin, 1877; Lewis, 1933; Dore, 1974; Bruner, 1975; Halliday, 1975).

In his analysis of young children's prosody, Furrow (1984) made psychophysical ratings of each sampled utterance's pitch, loudness and pitch range on a five point scale, as 'low', 'medium low', 'medium', 'medium high' or 'high'. Ratings on these three features were then aggregated into a single prosodic score by assigning numerical values to only three of these ratings (i.e. low=1, medium low=1.5, high=3), and then adding the three scores for each utterance.

Marwick (1986), who studied the intonation of vocalisations by two toddlers and their mothers, also employed an impressionistic system of analysis. Intonation in each syllable was assessed as 'high', 'medium high', 'medium low', or 'low' and described as 'level' or travelling 'upwards' or 'downwards'. This system was devised to described the dynamic movement of the intonation in each utterance that is, the pitch excursions in terms of height and range, and to provide at the same time details about the intonation shape before and after the pitch excursions, as well as the overall shape. This method springs directly from the linguistic tradition which sees prosody from the point of view of its place in the linguistic structure rather than its phonetic nature (Ladd and Cuttler, 1984).

Despite the fact that the methods of subjective dimensional ratings rely on what a hearer can perceive, they are not reliable, for they refer, intentionally or unintentionally, to a cultural and/or language norm, which serves as a baseline for comparison of each infant vocalisation. Moreover, Marwick's (1986) method employs the 'syllable' as the unit for the pitch movement, so it is not appropriate for the analysis of prosody in prelinguistic vocalisations where the form of syllable has not yet fully developed.

Some subjective methods of analysis of infant prosody transcribe vocalisations using the notation system of Western classical music. This approach depends on the assumption that prosodic features correspond to the elements involved in musical composition. Following the musicological approach Marcos (1987) asked a professional musician to first listen to each vocalisation and then produce it vocally, and define the tone to the nearest half-tone using an electronic tuner. In this way all the vocalisations were transcribed as notes of the semitone scale in the tone system of Western music. For pitch direction, vocalisations were placed into one of five categories: 'rising', 'falling-rising', 'level', 'falling', and 'rising-falling'. In the same vein, the Papouseks (1981) analysed their daughter's vocalisations from birth to 16 months with four procedures (a) musical transcription by two independent musically trained assistants (b) phonetic transcription with a simplified version of the International Phonetic Association (c) spectrographic analyses of sounds, and (d) automated analyses of temporal structures.

A musical notation system employs particular cultural conventions and serves to display a musical piece that has been composed to be repeated in a specific way by different performers, so as every time to convey to an audience the idea(s) and/or emotion(s) the composer had in mind. Thus, a musical notation system is not appropriate for a phenomenological description of spontaneous infant vocalisations, which are assumed to express a variety of messages in mother-infant interaction. In the notation system of the Western classical music, each note is discrete, having been given an absolute physical value, whereas the tones of speech are relative and produced in a continuum. In other words, the Western classical music notation system is a rigid system that cannot adequately represent the natural flow of a vocalisation. If one tries to represent the movement of an infant vocal utterance as a sequence of notes, one is forced to transcribe the changing pitch level as a sequence of notes that corresponds most closely to what is perceived, but which is not necessarily what is actually perceived.

4.3.3 The Method of Analysis of Infant Prosody followed in the Present Study

This study investigates the role of infant prosody, as an expressive system in its own right that is assumed to convey messages in mother-infant communication. According to this aim, an objective and not linguistically biased account of infant prosodic features needs to be obtained which would also somehow represent what the mother can perceive. A method for the analysis of prosody in infant vocalisations that would combine the advantages of both the acoustic measurements and the subjective assessments was felt to be the most appropriate.

As a first step, objective representation of prosodic features would be obtained through acoustic analysis. As Lynip (1951) has noted, by adopting an acoustic description it is possible to free oneself from the linguistically defined structure of phonetics, allowing one to describe the infant sounds in terms of parameters applicable to any vocalisation, regardless of the maturity of the person vocalising. Then these objective values would be transformed into a psychophysical scale, from which, in turn, several measurements describing the dynamic movement of the vocalisation would be calculated.

This approach relies on the assumption that the perceptual features of sound, i.e. pitch and loudness, depend exclusively on their acoustic correlates, that is, on fundamental frequency and intensity, respectively. However, the human auditory system does not perceive sound parameters in isolation, rather what is perceived is the product of the inter-relation between these parameters. This integrative mode of functioning of the auditory system is a major asset for speech perception (Scherer, 1982), but they impose limitations in any method of analysis of prosodic features.

4.3.3.ai Auditory Assessment of Pitch and Loudness

a) Pitch

Although pitch depends mainly on fundamental frequency, it is also determined by other factors such as intensity and duration. In particular, sounds of low frequency were found to fall in pitch with increasing intensity, whilst tones of high frequency rise in pitch with increasing intensity, and tones of middle frequency (1-2 kHz) show little change. However, it has been demonstrated that this effect is very small and does not substantially affect listeners' hearing (Small, 1973b; Rossing, 1989).

Regarding the parameter of duration, very brief tones are described as 'clicks', but as the tones lengthen, the clicks take on a weak sensation of pitch, increasing in strength upon further lengthening. For tones with frequency up to 1500 Hz the lower the frequency of the sound the longer the duration required for a given tone to produce a definite pitch. For example a tone of 500 Hz needs 20 milliseconds to produce an awareness of pitch (Rossing, 1989).

Despite all the above considerations, it is widely accepted in studies of adult and infant prosodic features that vocal pitch can be reliably represented by measurements of fundamental frequency.

Pitch Shift Discrimination

Pitch shift discrimination, i.e. the ability to hear difference in pitch, has received considerable attention in the literature on pitch characteristics of tonal stimuli. Most of the studies have employed experimental methods using as stimuli synthetic pure tones, or tones of 1000 Hz or more (Sergeant and Harris, 1962; Pollack, 1968; Tsumura,1976, 1985; Tsumura et al., 1973, 1990; Gulick, 1989). Rossi (1971) and Kaltt (1973) measured the absolute threshold of pitch change in speech-like signals. In his experiment Rossi (1971) used as stimuli glides produced by a male voice, starting at approximately 135 Hz and ending at various frequencies from 0 up to about 60 Hz higher than the onset frequency. Durations were 200, 100 and 50 milliseconds. According to the results, the threshold values are 95, 250 and 890 Hz/sec respectively. Klatt (1973), however, reported that a glide in a synthetic vowel /ɛ/ starting at approximately 120 Hz was audible if the rate of change was only 12 Hz/sec.

In an attempt to find a compromise between these diverging results, Lehiste (1970) suggests to convert the values for the threshold slopes into semitones per second

instead of Hz per second, in order to make them independent of the frequency region in which the experiments had been done. After this transformation it comes out that a short glide of 50 milliseconds is audibly distinguished from a stationary tone if the rate of change of its fundamental frequency is 64 semitones per second. Each doubling of the stimuli duration decreases the threshold value by a factor of four.

D'Odorico and Franco (1991) classified the melody type of the infant fundamental frequency contours as 'level', 'simple rising', 'simple falling', 'complex rising' and 'complex falling'. Only changes in pitch level equal to or greater than 47.8 Hz were considered. Flax et al. (1991) classified only the final contour of the fundamental frequency as rising, flat or fall. The contour was considered to rise or fall if the movement was greater than 40 cycles and lasted 150 milliseconds or longer. The authors adopted these criterion on the basis of earlier pilot data, where fifty vocalisations of a prelinguistic child, not included in the final study, were compared relative to perceptual judgements and acoustical analysis of contour direction. Perceptual judgements of these categories were only consistent with acoustic analysis when a change in direction was greater than 40 cycles and lasted more than 150 milliseconds.

One cannot be sure, however, that the results described above are generally valid in actual speech, where the sounds are not heard in isolation, as is the case in the experiments. A pilot study based on a sample of 1218 vocalisations collected in the Main and the Pilot study, which were judged by two experienced musicians, showed that the pitch shift discrimination depends on several factors among which the following are outstanding: intensity- the stronger a tone is the better the discrimination; duration- the most favourable discrimination occurs when there is an abrupt transition from one pitch to the next within a tone; timbre- discrimination varies with both degree and kind of richness, ensemble of stimulus condition and training of the listener, These observations are in agreement with studies on pitch shift discrimination of complex tones, including adult speech (Rosenblith and Stevens, 1953; Seashore, 1967; Lehiste, 1970; Rossing, 1989).

I would agree with Rosenblith and Stevens (1953) who doubt the wisdom of postulating a 'true' pitch difference limen, since it depends so greatly on a given set of conditions, and suggest that it would be more reliable pitch shifts to be judged subjectively.

b) Loudness

As was noted earlier, the physical correlate of loudness is intensity. However, loudness of vocal sounds also depends upon other factors such as the fundamental frequency, the spectral characteristics of the sound as well as on its duration.

The basis of the dependence of language upon frequency is the different sensitivity of the ear to different frequencies. At the lowest frequencies, a much greater intensity is required for a sound to be audible than at higher frequencies. The ear is most sensitive to frequencies between 1000 and 6000 Hz, where much of the crucial acoustic information for speech is found. Below 1000 Hz, loudness sensitivity drops off steeply with descending frequency (Lehiste, 1970; Laver, 1994).

The dependence of loudness of a complex sound such as human vocalisations upon its spectral characteristics results from the fact that loudness of complex tones is equal to the sum of the loudness of the several components. In particular, a factor that determines loudness is the interaction between fundamental frequency and formant frequency. If the articulatory configuration of the vocal tract remains fixed but the fundamental frequency of the voice is changed, extensive changes in the overall intensity level will occur in the sound heard. Specifically, the intensity will increase if a harmonic coincides with the frequency of one of the lower formants, especially the first formant- since most of the sound energy of the vowel is contained in the first formant (Lehiste, 1970).

Moreover, the loudness of a vocal sound is determined by the separation of formants, or the formant structure. In particular, when the formant separation exceeds a critical bandwidth, the total loudness begins to increase (Gulick, 1989; Rossing, 1989). It has also been shown that the formant structure does not remain constant throughout the vowel, the initial and terminal frequencies of the formants being determined by the preceding and following consonant (Lehiste, 1970).

The threshold of hearing is lowered by increasing the duration of the sound up to duration at least as long as one second. The loudness of an intense noise, however, depends upon its duration up to duration of only 50 milliseconds for a sensation level of 10 dB, decreasing to 15 milliseconds at 60 dB sensation level (Lehiste, 1970; Gulick, 1989; Rossing, 1989). In the interpretation of intensity curves the 'intrinsic sonority' of each segment of speech sounds must also be taken into consideration. This feature is defined as the segment's loudness relative to that of other segments with the same

length and pitch (Lehiste, 1970; Laver, 1994). In studies of these effects in real speech it was found that the differences between the maxima and minima of the intensity curve were of the order of several decibels (Ladefoged, 1962).

Finally, the intensity of a vocal signal depends on the speaker's orientation and distance in relation to the microphone. In the present study of spontaneous behaviour the babies were necessarily free to move in the room, so their distance from the microphone as well as the orientation of their heads relatively to the microphone was continuously changing.

In order to make meaningful comparisons between the intensity curves of vocal sounds one must control all the other factors described above except the vocal effort of the speaker. This task clearly goes beyond the possibilities of the present study and, for this reason, measurements in loudness were not taken. Nevertheless, informal observations of selected vocalisations, where the infant was stationary showed that loudness may play an important role in the communication of emotions and communicative functions.

4.3.3aii Timbre

It has been demonstrated that emotions, mental states and communicative functions are significantly associated not only with pitch and loudness but also with timbre (Scherer, 1986; Fernald, 1993). The term timbre is used to denote the 'tone quality' or 'tone colour' of a sound. The American National Standards Institute defines timbre as that attribute of auditory sensation in terms of which a listener can judge two sounds similarly presented and having the same loudness and pitch are dissimilar. Timbre depends primarily on the energy distribution in the frequency range and the number of harmonics (Rossing, 1989). Although timbre depends on the prosodic features of pitch and loudness, it does not itself considered as a prosodic feature (see Chapter 1). Timbre has usually been described in terms of pairs of opposite verbal attributes (von Bismark, 1974; Pratt and Doak, 1976). However, this subjective method has been criticised (Rossing, 1989) and it does not seem to be an agreement among the researchers upon a reliable method for the description of timbre.

4.3.3b Some Considerations for the Acoustic Analysis of Pitch

A factor that should be taken into account in any method of prosodic analysis is the existence of lower level interactions between articulatory formations and pitch. Effects of this kind mostly concerned fine-grained perturbations of pitch caused by the

muscular and aerodynamic features of momentary articulations (Lehiste, 1970; Ohala, 1978; Laver, 1994). Two main kinds of pitch perturbation have been observed:

a) Microperturbations of pitch: These are very small scale perturbations of laryngeal periodicity in the train of individual pulses in a phonatory sequence. They are caused by the fact that virtually every act of voicing does not show a perfectly regular repetition in its timing, rather it shows some minor dysperiodicity, which is perceived as a very slight roughness in the phonation rather than as a change in melody (Laver, 1994).

b) Microprosodic perturbatory effects on articulatory formations on pitch: When sets of vowels or vowel-like sounds are pronounced by a speaker in a standard context, measurements show that each of these is associated with different average pitch, the 'intrinsic pitch' (Laver, 1994). In particular, 'high' vowels such as [i] and [u] tend to have higher fundamental frequencies than 'low' vowels such as [a] and [æ]. It has been suggested that intrinsic pitch differences on vowels are likely to be more a language universal characteristic of the speech production system, as opposed to characteristics which are language dependent (Laver, 1994).

Whalen and colleagues (1995) looked for the intrinsic fundamental frequency characteristics in the babbling of 12 infants 6, 9, and 12 months old, 6 of each in English- and French-learning environments. According to their results, intrinsic fundamental frequency associated with vowels appears even in babbling; there was no evidence of a change in the effect across the three ages examined, and no differences were found between the two languages.

Although the microprosodic distortions physically imposed by adjustments of the vowel or consonant formations on the vibratory pattern of the vocal folds, are often substantial, it seems that the listener discounts such distortions when perceiving the pitch contour of the vocalisation (Lehiste and Peterson, 1961; Reinholt Petersen, 1986; Laver, 1994).

Experimental studies on the amount of the influence of the vowel quality on the fundamental frequency an utterance is produced, have shown that the maximum fundamental frequency difference between a 'high' and a 'low' vowel that are perceived at equal pitch is less than 3 Hz (Chuang and Wang, 1978; Reinholt Petersen, 1986). This difference does not exceed one tenth of a semitone.

4.3.3c Description of the Method for the Analysis of Infant Prosody

A preliminary examination of this corpus of vocalisations revealed that it does not form a homogeneous group, rather they show variety in duration (lasting from less than 100 milliseconds to almost 5 seconds), fundamental frequency (ranging approximately from 267 to 1500 Hz), external structure (being either single vocalisations or sequences) and external morphology (consisting either of a single resonant nucleus (vowel-like sound) or combinations of a resonant nucleus with a non-resonant nucleus (consonant-like sound). The external structure and the external morphology of a vocalisation do not themselves constitute prosodic features, but they constitute the basis for the modulations of the prosodic features. This diversity in characteristics should be taken into account in a method of analysis of prosody.

Seven hundred and thirteen vocalisations were analysed in this study; 322 for the boy, and 391 for the girl. A vocalisation was defined operationally as any egressive voiced sounds, i.e. any sound produced by the periodic vibrations of the vocal cords during exhalation.

The analysis of prosody was preceded by auditory descriptive classifications of the vocalisations according to their external structure and external morphology. The classifications were based on the following categories:

A) *External Structure*:

1) Single vocalisations

2) Sequences containing single vocalisations of the same or different types that characteristically appear in this form and that can be considered as comprising a single behavioural pattern, or one kind of utterance.

B) *External Morphology*

1) Simple vocalisations consisting of a resonant nucleus.

2) Complex vocalisations consisting of a resonant nucleus (vowel-like sound) combined with a non-resonant nucleus (consonant-like sound).

Sequences of single vocalisations which are composed by the combination of a 'vowel-like' sound with a 'consonant-like' sound are called 'babbling' (Oller, 1980). In the sequences of vocalisations the duration of each component and the pauses between them are measured as well. These classifications might reveal important developmental changes.

As a first step in the prosodic analysis of the data corpus, a quarter range, high resolution, narrow band (25 ms/40 Hz) spectrogram was generated for each vocalisation using the Signalyze™ software for sound analysis (Keller, 1994). The narrow-band spectrogram is more appropriate than a wide-band spectrogram for the purposes of this study because, while the latter is more suitable for analysing formants and articulatory aspects, the first yields measures of the harmonic structure of the voiced sounds. Moreover, the interpretation of a wide-band spectrogram is particularly difficult for voices with high fundamental frequency like those of infants (Scherer, 1982; D'Odorico and Franco, 1991). The spectrogram was used to get information on the amount of noise embedded in the vocalisations, in order to decide if reliable fundamental frequency curves could be extracted. It was also needed to measure the duration of the vocalisations.

Due to the nature of voice production, at the beginning and the end of each vocalisation there is an area where the noise to signal ratio is high. Fundamental frequency curves calculated with these areas include at the edges extraneous values which do not belong to the tonal movement of the vocalisation. Therefore, in order to obtain fundamental frequency curves reliably representing the tonal movement of the vocalisation, the 'noisy' areas should be specified by measuring the noise to signal ratio at the edges, and excluded from the calculations. However, the software used for the acoustic analysis in this study did not include a command for this measurement. As an alternative method, a fundamental frequency curve was calculated every one millisecond for each vocalisation via the Autocorrelation routine. This curve displays precisely enough the tonal movement of the vocalisation, breaking off at the beginning and the end of the vocalisation where the noise to signal ratio increases. Calculations made over a time window wider than one millisecond tend normalise the curve, especially at the edges, giving a false display of the pitch movement. So far, none of the researchers who have analysed pitch in infant vocalisations acoustically has attempted to control for this factor.

Fundamental frequency curves obtained every one millisecond are not appropriate, however, for obtaining further measurements on the pitch movement, because they contain superfluous details. Thus, for each and every vocalisations another fundamental frequency curve was calculated every five milliseconds and measurements were obtained from the part of this latter curve, which corresponds to the continuous part of the curve calculated every one millisecond, thus, leaving out any false values at the edges of the curve.

As was noted earlier, the acoustic feature of fundamental frequency and the perceptual feature of pitch are not linearly related. Pitch increases less and less rapidly as the stimulus frequency is increasing linearly, and more and more rapidly as the stimulus frequency is increased logarithmically (Scherer, 1982). Thus, in order to get measurements more sensitive to what the mother could perceive, the raw fundamental frequency values describing each vocalisation were converted into values on a semitone scale. The semitone scale is a 12-note scale consisting of equally spaced intervals which correspond to equal frequency ratios, regardless of the absolute values of frequency. The size of an interval on a semitone scale is proportional to the logarithm of a frequency ratio i.e., pitch interval (in semitones)= $12\log_2 (f2/f1)$. The logarithm is to base 2 because the fundamental frequencies of two notes at the edges of a frequency scale (octave) differ by the factor 2. The value of each note on a semitone scale is given by the formula $12\log_2 (F_0)$ (Campbell and Greated, 1987; Roederer, 1995)[1] . While this is not the only way to represent the relation between frequency and perceptual response (Lynip, 1951), it is conventionally accepted as an appropriate means of normalisation, and has been used in analysis of prosodic characteristics in maternal speech to infants (Grieser and Kuhl, 1988; Fernald et al., 1989).

The following measurements were calculated from the pitch curves obtained as described above:

(1) Highest pitch
(2) Lowest pitch
(3) Initial pitch
(4) Final pitch
(5) Ratio of Initial and Final pitches
(6) Pitch Range
(7) Mean pitch
(8) Standard deviation of the mean pitch
(9) Duration of the vocalisation calculated from the spectrogram
(10) Duration from the Initial to the Highest pitch.
(11) Duration from the Initial to the Lowest pitch.
(12) Pitch at the highest point of the steepest rising slope in an undulating vocalisation.
(13) Duration from the Initial to the highest pitch point of the steepest rising slope in an undulating vocalisation.

---

[1] For describing changes in intonation in adult speech the basic interval of the semitone scale is often subdivided into 100 equal intervals. The unit of this subdivision is called *cent* and the pitch interval in cents is given by the formula: $1200\log_2 (f2/f1)$.

(14) Average Rate of Pitch Change.
This parameter was calculated as the ratio of the pitch range to the duration of the change. For bell shaped contours, the average rate of change was calculated as the ratio of the difference between the peak pitch and the mean pitch of the beginning and the end pitch to the duration of the curve. In the case of undulating contours, the rate of change of the sharpest bell was calculated. In the last two cases the average rate of change of the steepest rise slope was also calculated .

(15) Pitch Contour Shape:
Pitch movement was judged on a subjective basis; the researcher's musical training can be considered to approximate the mothers' sensitivity to their infants' vocalisations. The reliability of these subjective judgements was checked by asking a second experienced musician to listen to part of the corpus and classify it. The percentage of agreement between the two judges was 82%. Six categories were used for the classification of the direction of the pitch movement: 'level', 'rise', 'fall', 'bell', 'undulating rise' and 'undulating fall'.

In unit vocalisations measurements (1), (2), (3), (4), (5), (6), (7) and (8) were taken for the whole utterance. The rate of change is not calculated.

The method described above is applicable not only to human vocalisations, both infant and adult, but also to those of other species. It is believed that such a non-linguistic ethological description of the prosodic features in infant vocalisations will contribute to understanding of both the ontogeny and the phylogeny of language.

4.4 Statistical Analysis

The data collected at 2 week intervals were collapsed every two recording sessions, and the results are presented for every month. For each age an Analysis of Variance (ANOVA) was carried out to compare the means of the distributions obtained from measurements of each and every prosodic feature of the vocalisations included in the different message categories reported by the mother. The ANOVA tests compare the means of distributions, taking into account the variation *within* each distribution as well as the variation *between* the distributions. One-Way ANOVA was preferred to Two-Way ANOVA, which compares Message x Age x Prosodic Feature, since, as a consequence of the infants' psychological development, the same messages did not appear at every age.

Making several comparisons among a group of means produces a high probability of making a type I error, i.e. rejecting the null hypothesis that the means do not differ when different effects, in fact, occur. This is called Family-Wise Error Rate. Many multiple-comparison procedures have been devised for controlling the Family-Wise Error Rate. In this study $Tukey_a$ test, also called the Honestly Significant Difference (HSD) test, was used. The Tukey HSD is generally regarded as the best procedure for controlling the probability of a Family-Wise Error when all pairwise comparisons are made, because the difference between any two means is compared to the maximum difference of the means (Howell, 1992; Coolican, 1994).

The parametric assumption that has to be satisfied before proceeding with One-Way ANOVA test is the *homogeneity of variance*. This condition is satisfied when the variances of the distributions whose means are compared are not significantly different. In order to meet this requirement, the values of all the distributions compared were transformed logarithmically prior to the test. A distribution was included in the statistical analysis if it included four or more observations.

Given that the values of the expected frequencies were not very small, the association between the variable of 'message' and the variable of 'contour shape', which are unrelated data at the nominal level, was measured by the Chi-Squared test (χ^2). The statistical package MINITAB used for the analysis automatically checked, in each case, if the condition of not very small values of the expected frequencies was satisfied.

If vocalisations belonging in different message categories do not differ significantly in any of the prosodic features measured, these vocalisations are supposed to share the same prosodic form. A prosodic form will henceforth be called 'emphatic' when it has comparatively higher rate of pitch change. The term 'alerting' will be characterising forms which show significantly higher values in measurements of 'pitch height' (i.e. 'peak pitch', 'mean pitch', 'beginning pitch', 'final pitch') or in the 'rate of pitch change' than the form(s) that is compared with.

For the messages or similar prosodic patterns, a linearity test was used to detect any modifications with age that appeared over a period of three or more months.

CHAPTER 5

RESULTS: ROBIN, JULIE

5.1 Introduction

The results will first place be presented for each subject separately, followed by a comparison of the two subjects.

In the tables presenting the prosodic patterns of different messages and their characteristics, the first column next to each prosodic feature represents from which of the message(s) differs the darkest shadowed message(s) and the ANOVA results. The second column represents differences from the lighter shadowed message(s). In the boxes are presented the median and the range of each distribution. These are raw values, i.e. before the logarithmic transformation (Appendix II).

5.2a Robin

7 MONTHS (N = 72)

a) Messages attributed by the Mother

At seven months Robin's mother reported messages classified as 'assertive' and 'emotional'. The 'assertive' messages were either 'interpersonal', namely 'Attention Seeking' and 'Invitation to Stranger', or 'referential', namely 'Non-Shared Interest' and 'Shared Interest'. At this age three of the vocalisations in the 'Request' category were recognised as 'Requests for Action on Object' and one was interpreted as 'Request for Action by the Mother herself'. The emotional message reported was 'Shared Negative Emotion'.

b) Analysis of Prosody

1) Contour Shape

Regarding the final movement of the contour shape, vocalisations classified as 'Attention Seeking', 'Invitation to Stranger', 'Shared Interest', and 'Non-Shared Interest', compared as a group with vocal expressions of 'Shared Negative Emotion' and 'Requests', display contour shapes with rising end significantly more often than the vocalisations in the latter two categories, which are equally distributed in patterns with rising and non-rising end ($\chi^2 = 10.44$, df = 2, $p<.01$). The utterances in the first group (39) are mainly rises, whilst the majority of 'Shared Negative Emotion' vocalisations are either undulating rises or undulating falls ($\chi^2 = 6.34$, df = 1, $p<.02$) (Table 5.1).

2) Prosodic Forms

From the analysis of prosody three prosodic forms were extracted on the basis of significant differences in 'peak pitch', 'mean pitch', 'duration', 'pitch range', and 'rate of pitch change' (Table 5.2).

Form A: Can be described as relatively short, low pitched, narrow ranged, with low rate of pitch change. This form is displayed by vocalisations interpreted as conveying 'Invitation to Stranger', 'Non-Shared Interest' and 'Shared Interest'.

Form B: Characterises 'Attention Seeking' vocalisations and differs from Form A and Form C, described below, in that it is more emphatic.

Form C: Is longer than both Pattern A and Pattern B and has higher pitch level and wider pitch range than Pattern A. This pattern expresses 'Shared Negative Emotion'.

Vocalisations classified as 'Requests' do not display a specific prosodic form, rather they share characteristics with all the other forms. In particular, vocal 'Requests' are significantly longer utterances than forms A and B. Regarding the parameters of peak pitch, mean pitch, pitch range and rate of pitch change, however, 'Requests' do not differ significantly from either forms A, B or C.

3) Conclusion

The results described above demonstrate that at seven months vocalisations interpreted as 'Shared Interest' and 'Non-Shared Interest' are conveyed by the same prosodic form, whilst vocalisations recognised as expressing 'Attention Seeking' and 'Shared Negative Emotion' show each a distinct prosodic form. Thus, it seems that for this infant at this age prosody does not distinguish all the kinds of messages that the mother perceives; rather, three groups of messages are differentiated by the way they are voiced, namely 'referential' messages, 'interpersonal' messages, and 'emotional' messages. Exception to this three-way classification are vocal 'Invitations to Stranger'; although this is an 'interpersonal' message, it is conveyed with the same prosodic form as the 'referential' vocalisations. At this age this boy uses a distinct prosodic form only for interpersonal messages addressed to the mother that is, for the message 'Attention Seeking'.

c) Video Analysis

The video analysis demonstrated that in the message groups identified by analysis of prosody, individual messages are distinguished by the mother on the basis of the accompanying non vocal behaviour as well as the timing of the vocalisations[1] . In particular, for the utterances interpreted as 'Non-Shared Interest', 'Shared Interest' and 'Invitation to Stranger', it was found that the accompanying non-vocal behaviour, functionally classified from the video corresponds to the message attributed by the mother (Table 5.3).

[1] It should be noted that for the distinction of individual messages the mother may have also relied on other vocal characteristics, such as loudness and timbre which, however, are not studied in the present thesis.

In this corpus two kinds of 'Requests' were distinguished, namely 'Requests for Action on Object' (3) and 'Requests for Action by the Mother herself' (1). As Table 5.3 shows, the non vocal behaviour accompanying these utterances implies mainly a motivational state of 'Converging Interest' or 'Response to the Mother's Invitation to Play' respectively. None of the 'Requests' is combined with an imperative gesture (see Appendix I). However, the timing of these vocalisations may be the cue that the mother uses to attribute this message. Specifically, 'Requests for Action on Object' always follow a change of activity in an 'active object', and 'Request for Action by the Mother herself' occurs during a body game immediately after a pause in the mother's rhythmic activity.

'Attention Seeking' is, in most cases, conveyed prosodically, as well as through the accompanying non-vocal behaviour (Table 5.3). In one occasion the mother reported that the infant vocalised to invite her for an interpersonal exchange while she was ignoring him. However, according to the video analysis, the behaviour accompanying this vocalisation was object inspection. A closer examination of the prosodic features revealed that this vocalisation did not deviate from those which characterise utterances belonging to the 'Attention Seeking' category. The mother was apparently highly sensitive to the vocal expression of her infant.

Vocalisations recognised by the mother as expressing 'Shared Negative Emotion' were uttered in a variety of situations, namely 'Converging Interest', 'Conflict' and 'Response to Invitation to Play' (Table 5.3). At seven months this infant expresses 'Shared Negative Emotion' by means of a prosodic form but also through facial expression of frustration, which accompanied all the vocalisations of this type.

8 MONTHS (N = 53)

a) Messages attributed by the Mother

As was the case at seven months, at eight months Robin's mother reported messages classified as 'assertive' and 'emotional'. Some of the 'assertive' messages were 'interpersonal', namely 'Invitation to Stranger', and some 'referential', namely 'Shared Interest'. The messages 'Attention Seeking' and 'Non-Shared Interest' that were reported at seven months, do not appear at this age. All the 'Requests' at eight months were 'Requests for Action on Object'. In the class of 'emotional' messages except the 'Shared Negative Emotion' category now also appears the category 'Shared Positive Emotion'.

b) Analysis of Prosody

1) Contour Shape

Vocalisations recognised by the mother as conveying 'Invitation to Stranger', 'Shared Interest', 'Request' and 'Shared Negative Emotion' are related to contour shapes with rising tail, while the vocal expressions of 'Shared Positive Emotion' are equally distributed in patterns with rising and falling end ($\chi^2 = 4.9$, df = 1, $p<.05$). Regarding the overall contour shape, 'Invitation to Stranger', 'Shared Interest', 'Request' vocalisations compared as a group with vocal expressions of 'Shared Negative Emotion' are mainly 'rises' whereas the latter are 'undulating rises' ($\chi^2 = 23.63$, df = 1, $p<.001$). Vocalisations conveying 'Shared Positive Emotion' are mainly bells or rises (Table 5.4).

2) Prosodic Forms

At eight months vocalisations conveying different messages were statistically differentiated in the parameters of 'peak pitch', 'mean pitch', 'beginning pitch', 'final pitch', and 'duration'. From these results the following prosodic forms emerged (Table 5.5):

Form A: Corresponds to Form A at seven months, and it can be described as short and low pitched. 'Invitations to Stranger' are again conveyed by this form.

Form B: Is more high pitched than Form A and Form C described below. This form characterises vocalisations interpreted by the mother as 'Requests'.

Vocalisations recognised as conveying 'Shared Interest' are not statistically differentiated from vocalisations displaying either Form A or Form B in any of the prosodic features measured.

Form C: Expresses 'Shared Negative Emotion'. It is the longest of all the other forms that appeared at eight months. Also, Form C is more high pitched than Form A.

Form D: Consists of a very high pitched register compared to all the other forms and characterises 'Shared Positive Emotion' vocalisations.

3) Conclusion

Similarly to seven months, prosody at eight months distinguishes the 'referential' messages from the 'emotional' messages. Moreover, the vocalisations expressing 'Shared Negative Emotion' and 'Shared Positive Emotion' have different prosodic forms.

c) Video Analysis

The video analysis demonstrated that, although the vocalisations conveying 'referential' messages do not display different prosodic forms, they are attributed different messages on the basis of the accompanying non vocal behaviour and the timing of their occurrence.

In particular, for the utterances interpreted by the mother as 'Shared Interest' it was found that, in most cases (71%), the accompanying non-vocal behaviour conveys the message attributed by the mother (Table 5.6). As Table 5.6 shows vocalisations conveying 'Shared Interest' and 'Request for Action on Object' are in most cases accompanied by the same non-vocal behaviour which conveys 'Converging Interest'. Nevertheless, most of the vocal 'Requests for Action on Object' (8 or 67%), unlike utterances interpreted as 'Shared Interest', immediately followed a change in an 'active' object. In the remaining 4 occasions for the interpretation of the vocalisations as 'Requests for Action on Object', the mother took into account the infant's knowledge about the objects' possibilities. Explaining her answer about these vocalisations the mother reported that the infant could remember

that a particular object (i.e. a music box) could be activated to produce an interesting effect. Thus, she assumed the infant could remember such an effect in that object.

Regarding the emotional vocalisations, the majority of utterances of 'Shared Negative Emotion' occurred in situations of potentially negative character that is, 'Conflict' (10) and 'Attention Seeking' (3). On 3 occasions 'Shared Negative Emotion' was expressed in situations of 'Converging Interest', after a change of an 'active' object, a situation where, on other occasions, vocalisations interpreted as 'Requests for Action on Object' have been found. Vocalisations expressing 'Shared Positive Emotion', on the other hand, are accompanied by the same non-vocal behaviours as the 'Shared Interest' utterances. The expressions of 'Shared Positive Emotion' and 'Shared Negative Emotion' are mainly distinguished by their prosodic forms. Moreover, 67% of the 'Shared Positive Emotion' vocalisations are combined with a smile, whilst this is the case only for 1 of the 7 vocal expressions of 'Shared Interest'. Only half (8) of the vocalisations expressing 'Shared Negative Emotion' are accompanied by a facial expression of frustration.

For the vocalisations classified as 'Invitation to Stranger' the functional analysis of the accompanying non-vocal behaviour recorded on video correlates with the mother's interpretation (Table 5.6).

9 MONTHS (N = 74)

a) Messages attributed by the Mother

At nine months appear 'assertive' and 'emotional' messages, as in seven and eight months, but also for the first time the mother reports 'receptive' messages. The 'assertive' messages were also classified as 'interpersonal', namely 'Attention Seeking' and 'Invitation to Stranger', or 'referential', namely 'Non-Shared Interest' and 'Shared Interest'. At this age the mother identified two kinds of 'Requests', specifically 'Requests for Action by the Mother herself' (6) and 'Requests for Action on Object' (2). The 'receptive' messages reported were 'Shared Comment' and 'Shared Conversation'. The 'emotional' messages identified were 'Shared Positive Emotion' and 'Shared Negative Emotion'.

b) Analysis of Prosody

1) Contour Shape

The results show that contour shapes with rising end are no longer dominant, as they were at seven and eight months, and a richer variety of pitch contours is displayed. Table 5.7 shows the percentage distribution of the vocalisations of different message categories by contour shape. In particular, vocalisations conveying 'Invitation to Stranger', 'Attention Seeking', 'Non-Shared Interest', 'Shared Interest' and 'Shared Negative Emotion' are significantly more often related to contours with rising tail compared as a group with 'Shared Narratives', 'Shared Comments', 'Requests' and 'Shared Positive Emotion' utterances, which are significantly more often related to contours with falling tail ($\chi^2 = 12.78$, df = 2, $p<.01$). A closer examination of the vocalisations in each group reveals that vocal expressions of 'Invitation to Stranger', 'Shared Interest' and 'Non-Shared Interest' are mainly 'rises', vocalisations expressing 'Shared Negative Emotion' are mostly 'undulating rises', 'Shared Comments' are in the majority 'falls', and 'Requests' are 'bells'. 'Shared Narratives' and vocal expressions of 'Shared Positive Emotion' and 'Attention Seeking' are not related to any particular contour shape (Table 6.7).

2) Prosodic Forms

The analysis of prosody revealed three prosodic forms on the basis of significant differences in 'mean pitch', 'final pitch', 'duration' and 'contour shape' (Table 5.8).

Form A: Is short and low pitched, as Forms A at seven and eight months, but as opposed to those displays pitch contours with falling tail. This form conveys 'Shared Comments'.

Form B: Is longer than Form A and characterises 'Shared Narratives'.

Form C: Differs from both forms A and B in being more high pitched and displaying pitch contours with rising tail. This form conveys 'Shared Negative Emotion'.

Vocalisations conveying 'Invitation to Stranger', 'Attention Seeking', Non-Shared Interest', and 'Shared Interest' are more high pitched than the 'Shared Comments' (Form A) and shorter than 'Requests' and 'Shared Narratives' (Form B). Moreover, these vocalisations have contour shapes with rising end as opposed to forms A and B which display pitch contours with falling tail. In contrast with what happened at seven and eight months, at nine months vocalisations conveying 'Invitation to Stranger', 'Attention Seeking', Non-Shared Interest', and 'Shared Interest' do not differ prosodically from the vocalisations expressing 'Shared Negative Emotion'.

At this age vocalisations expressing 'Shared Positive Emotion' have lower pitch than the vocal expressions of 'Shared Negative Emotion' and the 'Invitations to Stranger'.

3) Conclusion

At nine months prosody no longer differentiates among 'referential', 'interpersonal' and 'emotional' classes of messages, as it was the case at seven and eight months. A distinction of a different character is revealed. In particular prosody now distinguishes between 'assertive' messages (i.e. 'Invitation to Stranger', Attention Seeking', 'Non-Shared Interest', 'Shared Interest') and 'receptive' ones (i.e. 'Shared Narrative' and 'Shared Comment'); the 'assertive' messages are expressed by

more alerting forms (i.e. more high pitched) than the 'receptive' messages. The vocal expressions of 'Shared Negative Emotion' are not prosodically differentiated from the vocalisations included in the 'assertive' group. The message 'Shared Negative Emotion' does not only express a certain feeling, it also has an 'assertive character, since it signals somebody's need for help. An exception to the distinction described above constitutes the 'assertive' message 'Request' which is expressed by a 'non-alerting' form.

c) Video Analysis

The message categories 'Invitation to Stranger', 'Attention Seeking', 'Shared Interest', and 'Non-Shared Interest' included in the assertive group, can be distinguished on the basis of accompanying non-vocal behaviour, which, in each case is in agreement with the message attributed by the mother (Table 5.9). Of the vocalisations expressing 'Shared Negative Emotion' only one was uttered in a 'Conflict' situation, while the rest occurred in situations where vocalisations conveying 'Attention Seeking' and 'Shared Interest' were found. In those cases the cue for the mother's attributions was probably the facial expression of frustration which accompanied all the vocal expressions of 'Shared Negative Emotion' at nine months.

As shown in table 5.9 only one vocalisation recognised by the mother as 'Request for Action by the Mother herself' was found, in the video analysis, to be combined with a gesture conveying the same message. The remaining vocalisations interpreted as 'Requests' are combined with the same non-vocal behaviours as 'Shared Conversations'. However, the characteristic difference between the situations of occurrence of these two kinds of vocalisations is the marking timing of the 'Request' vocalisations in the communicative episode. In particular, vocalisations recognised as 'Requests for Action by the Mother herself' were uttered immediately after a pause in the mother's rhythmic activity during a body game, or in a game situation just before the mother performed the action that she would be expected to, given the infant's knowledge about the sequence of the game. In the case of 'Requests for Action on Object' the vocalisations immediately followed either a change of an 'active' object, or a pause in the mother's activity on an object, after she had created an interesting stimulus for the baby to watch.

The majority of 'Shared Comments' (80%) occurred in situations of 'Converging Interest'. Finally, vocalisations interpreted as expressing 'Shared

Positive Emotion' occur in a variety of situations and all of them are combined with a smile (Table 5.9).

10 MONTHS (N = 68)

a) Messages attributed by the Mother

As was the case at nine months, at ten months Robin's mother reported 'assertive', 'receptive' and 'emotional' messages. The 'assertive' messages were either 'interpersonal', namely 'Attention Seeking' or 'referential'; this latter category included expressions of 'Shared Interest'. The five 'Requests' at this age were all for actions by the mother on an object. The 'receptive' messages were 'Shared Comment' and 'Shared Recognition of Object'. As at seven and eight months, the 'emotional' messages were 'Shared Positive Emotion' and 'Shared Negative Emotion'. However, for the first time at this age in the message category 'Shared Positive Emotion' the mother systematically distinguished two categories of vocalisations describing them with the semantically differentiated terms 'Enjoyment' and 'Excitement' (Plutchik, 1981). This distinction was kept in the analysis of the mother's interview.

b) Analysis of Prosody

1) Contour Shape

The change in the contour patterning observed since nine months proceeds further. Now the dominant tail shape of pitch is a fall (84%). As for the overall contour of the vocalisations, it was found that those conveying 'Shared Recognition of Object', 'Request' and 'Attention Seeking' are mainly bell shaped, whereas vocal expressions of 'Shared Negative Emotion' have mainly falling contours ($\chi^2 = 29.2$, df = 4, $p<.001$). Vocal expressions of 'Enjoyment' and 'Excitement' are, in the majority of cases, 'undulating falls' or 'bells'. Finally, 'Shared Comments' and utterances conveying 'Shared Interest' are not characterised by any particular contour shape (Table 5.10).

2) Prosodic Forms

The results show that at ten months only vocalisations of 'Shared Negative Emotion' show a prosodic form distinguished by high pitch. Otherwise, the prosodic features do not distinguish between main groups of messages, but do differ between some of the individual message categories. Thus, 'Shared Comments' proved to be significantly shorter than vocalisations of 'Excitement' and 'Enjoyment'.

Furthermore, vocal expressions of 'Excitement' have significantly wider pitch range than either 'Shared Comments' and utterances conveying 'Enjoyment'. Finally, 'Attention Seeking', 'Request' and 'Excitement' vocalisations are more emphatic than those expressing 'Shared Recognition of Object'. No reliable conclusions can be drawn for the vocalisations of 'Shared Interest' that appeared at this age, because of the great variability of the distribution of these expressions even after the logarithmic transformations. The variability of the 'Shared Interest' vocalisations in pitch level may express subtleties of motivation or interest (Table 5.11).

3) Conclusion

Although no consistent prosodic forms emerged, the same tendency as was observed at nine months continues; that is, 'Assertive' messages, such as 'Attention Seeking', 'Request' and 'Excitement' display 'alerting' prosodic characteristics (i.e. comparatively high rate of pitch change) compared to the more 'receptive' messages of 'Shared Recognition of Object' and 'Enjoyment'.

c) Video Analysis

At ten months interesting developmental changes occurred in situations that provoked 'Requests' and expressions of 'Shared Interest'. Regarding the 'Requests', in two cases the mother relied, as did in previous months, on the timing of the vocalisations; that is, the vocalisations followed immediately on a pause in the mother's activity on an object, after she had created an interesting stimulus for the baby to watch. However, for the first time the mother's interpretation of 'Requests' depended in three instances on her knowledge of the infant's past experiences of interactive play routines in 'Cooperative Game' and 'Book Reading'.

Cooperative games with an infant (e.g. putting wooden people in a toy car) rely on three requirements: (a) A shared plan for joint action to be performed immediately. This plan can be communicated by the mother or by the infant, (b) Each partner clearly attends to, and acts with reference to, the indicated interests and intentions of the other in the shared plan, or directs the partner's behaviour towards the completion of the shared purpose, and (c) Both partners willingly participate in the interaction by their own impulse and choice, without any coercion (Hubley, 1986). 'Book Reading' is built on already established skills such as recall, imposition of roles, turn-taking, joint attention and sequential structure; additionally 'Book Reading' involves an exchange about non-concrete, pictured topics (Bruner, 1983).

Once Robin reaches the stage of Secondary Intersubjectivity, the form of non-vocal behaviours suggesting 'Shared Interest' has also changed, and now, instead of simply conveying 'converging interest', it affirms 'joint interest'. These new behaviours and play routines reflect the infant's increasing ability to coordinate his interest and awareness with the mother's interest in objects, as is typical of the Secondary Intersubjectivity phase, described in Chapter 2 (Table 5.12).

With Robin the message 'Shared Recognition of Object' is reported by the mother for the first time at ten months. As was noted in Chapter 4, this message cannot be identified from the video record alone. According to the mother's report, her interpretation required recollection of Robin's previous experience with objects and the routines he has developed with those. The majority of the vocalisations interpreted as 'Shared Recognition of Object' were uttered in situations typical of the Secondary Intersubjectivity period, namely when the infant was signalling 'Joint Interest' or a 'Directive' (Table 5.12).

As was noted earlier, although vocalisations of 'Enjoyment' and 'Excitement' were distinguished prosodically from each other, they were not different from utterances classified in certain other message categories. In particular, 'Excitement' utterances were not prosodically different from those interpreted as 'Requests'. Also, vocal expressions of 'Enjoyment' were not prosodically distinguished from either the vocal 'Shared Recognitions of Object' or the 'Requests'. Furthermore, as Table 5.12 shows, 52% (14 vocalisations) of the 'Excitement' and 'Enjoyment' vocalisations as a whole is combined with gestures conveying 'Converging Interest', 'Joint Interest' or 'Directive', as the vocal 'Shared Recognitions of Object' and the 'Requests' do. As a whole, 96% of the 'Shared Positive Emotion' vocalisations were made with smiles. However, all the 'Shared Recognition of Object' utterances and 60% of the 'Requests' were also accompanied by smiles. In distinguishing between vocalisations expressing 'Excitement'-and-'Request', 'Enjoyment-and-'Shared Recognition of Object', and 'Enjoyment'-and-'Request', the mother may have relied on other aspects of the situation, past or present; namely, on her recollection of the baby's previous experiences with objects for utterances of 'Shared Recognition of Object', and on the timing of the vocalisations in the interactive play routines of 'Cooperative Games' and 'Book Reading' for vocal 'Requests'.

Vocalisations recognised as expressing 'Shared Negative Emotion' occurred in situations of 'Conflict', 'Attention Seeking' as well as when there was 'Converging Interest' after a change of an 'active' object where in other occasions 'Request for

Action on Object' have been identified (Table 5.12). At ten months Robin conveys 'Shared Negative Emotion' by means of prosody; the facial expression does not contribute to the expression of this emotion as decisively as before, since only half of these utterances are now combined with a facial expression of frustration.

For the vocalisations classified as 'Attention Seeking' the video analysis is in agreement with the mother's attribution (Table 5.12).

11 MONTHS (N = 59)

a) Messages attributed by the Mother

Similarly to the previous ages studied, at eleven months were reported the 'assertive interpersonal' messages 'Invitation to Stranger' and 'Attention Seeking', the 'assertive referential' message 'Shared Interest', and the 'emotional' messages 'Shared Positive Emotion' and 'Shared Negative Emotion'. For the first time, however, the mother identified ten vocalisations as 'Practising', which is judged as 'receptive' behaviour.

b) Analysis of Prosody

1) Contour Shape

Vocalisations interpreted as 'Attention Seeking', 'Invitation to Stranger' or 'Shared Interest' displayed in most cases a rising end compared to vocalisations recognised as expressing 'Shared Positive Emotion', 'Shared Negative Emotion' as well as those perceived as 'Practising', which were mainly characterised by contour shapes with falling end ($\chi^2 = 25.48$, df = 1, $p<.001$). As for the overall movement of the pitch contour, vocalisations conveying 'Attention Seeking', 'Invitation to Stranger' and 'Shared Interest' are mainly 'rises'. Vocal expressions of 'Shared Positive Emotion' are in the majority 'bells' and the 'Practising' utterances 'undulating falls'. Finally, 'Shared Negative Emotion' vocalisations are not related to any particular contour shape (Table 5.13).

2) Prosodic Forms

Three prosodic forms were extracted on the basis of significant differences in 'peak pitch', 'mean pitch', 'final pitch', 'duration', 'pitch range', 'standard deviation of pitch' and 'rate of pitch change' (Table 5.14).

Form A: Is short, low pitched, narrow ranged, with low standard deviation of pitch and rising contour tail. Form A conveys 'Attention Seeking' and 'Shared Interest'.

Form B: As opposed to form A, Pattern B is long, high pitched, wide ranged and with high standard deviation of pitch. This form expresses 'Shared Negative Emotion'.

Form C: Is displayed by vocalisations perceived as 'Practising'. It is long and more high pitched than Form A, but low pitched compared to Form B; it has wide pitch range, high standard deviation of pitch and falling contour tail.

At this age vocal 'Invitations to Stranger' are more high pitched than 'Attention Seeking' and 'Practising' vocalisations and also have higher standard deviation than utterances of 'Attention Seeking'. However, they do not differ prosodically from the vocalisations conveying any other message. Vocalisations expressing 'Shared Positive Emotion' are more high pitched than the vocalisations displaying Form A ('Attention Seeking' and 'Shared Interest') and Form C ('Practising'), but they are shorter utterances than those of Form C. Moreover, they have significantly wider pitch range and higher standard deviation of pitch than vocalisations of 'Attention Seeking'.

3) Conclusion

At eleven months vocal expressions of 'Shared Negative Emotion' display a register with comparatively the high pitch level. Also relatively high pitched are the utterances expressing 'Shared Positive Emotion'. The 'assertive' messages 'Attention Seeking', 'Invitation to Stranger' and 'Shared Interest' have contours shapes with 'alerting' rising end, whilst the 'receptive' vocal behaviour 'Practising' displays pitch contours with 'non-alerting' falling end.

c) Video Analysis

The analysis of the video record showed that the non-vocal behaviour accompanying utterances interpreted by the mother as 'Attention Seeking', 'Shared Interest' and 'Invitation to Stranger' correlated with the mother's attribution (Table 5.15).

At eleven months Robin's vocal expressions of positive and negative emotions are not prosodically distinguished. However, a close examination of the context from the video reveals that these are markedly different and that they do distinguish the infant's emotional state. Most of the vocalisations recognised as expressing 'Shared Negative Emotion' occur in situations of a potentially negative character; i.e. 'Conflict', 'Attention Seeking' or in situations of 'Converging Interest' immediately after an 'active' object has changed, where, in previous sessions, 'Requests for Action on Object' were found. In contrast, vocal expressions of 'Shared Positive Emotion' occur in situations of a potentially positive character; that is, when the infant

responds to the mother's invitation to play, when he himself invites the mother to play, or in situations of 'Joint Interest' (Table 5.15). Unlike what happened at seven, eight and nine months, at this age vocal expressions of 'Shared Positive Emotion' are not combined with a smile more often than the other non-emotional messages. Likewise, 'Shared Negative Emotion' utterances were not found to be accompanied by facial expression of frustration.

With regard to 'Practising' vocalisations, not only do these show a distinct prosodic form, but they are also the only vocalisations at eleven months that can be described as 'babbling' or speech-like sounds. One is struck by the fact that although these vocalisations occurred in a variety of communicative situations, mostly when the infant invites the mother for play (70%) (Table 5.15), the mother does not feel that they convey any particular message, rather she perceives them as 'vocal practice'.

Summary of Robin's Vocalisations

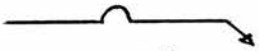
The boy's vocalisations from seven to eleven months displayed the following prosodic forms:

A. Non-Alerting

1) Short, low-pitched, monotonous, non-emphatic, with falling contour tail.

2) Medium length, low-pitched, monotonous, non-emphatic, with falling contour tail.

3) Long, low-pitched, monotonous, non-emphatic, with falling contour tail.



4) Long, low-pitched, fluctuating, non-emphatic, with falling contour tail.

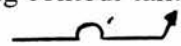
B. Alerting

5) Short, low-pitched, monotonous, non-emphatic, with rising contour tail.

6) Short, high-pitched, monotonous, emphatic, with rising contour tail.

7) Medium length, high-pitched, monotonous, non-emphatic, with rising contour tail.

8) Medium length, low-pitched, monotonous, emphatic, with rising contour tail.

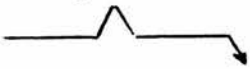


9) Medium length, high-pitched, monotonous, non-emphatic, with falling contour tail.

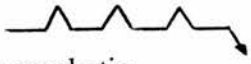
10) Medium length, high-pitched, fluctuating, non-emphatic, with falling contour tail.

11) Long, high-pitched, monotonous, non-emphatic, with rising contour tail.

12) Long, high-pitched, monotonous, non-emphatic, with falling contour tail.



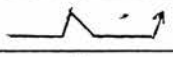
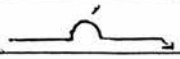
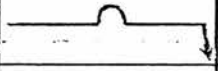
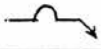
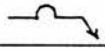
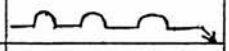
13) Long, high-pitched, fluctuating, non-emphatic, with falling contour tail.



14) Medium length, very high-pitched, monotonous, non-emphatic.

Table 5.16 shows the distribution of Robin's prosodic forms by age and class of messages. One can observe that before nine months prosody distinguishes between 'referential', 'interpersonal' and 'emotional' messages. The 'interpersonal' messages are expressed by more 'alerting' prosodic forms than the 'referential' messages. After nine months, which is the threshold to Secondary Intersubjectivity, the role of prosody changes. Prosodic forms in Robin's vocalisations now differentiate between 'assertive' and 'receptive' messages, the former being expressed by more 'alerting' forms than the latter. This distinction reflects a developing Self-Other awareness, which lays the groundwork for the development of the ability to coordinate one's own interest in an object with the mother's.

Table 5.16: Distribution of Robin's Prosodic Patterns by Age and Class of Messages.

MESSAGE	AGE (months)				
	7	8	9	10	11
Assertive					
Interpersonal					
Referential					
Self - Directed					
Other - Directed*					
*Requests**					
Receptive					
Interpersonal					
Referential					
Self - Directed					
Other - Directed					
Shared Narratives					
*Shared Comments**					
*Responses**					
Non-Shared Narratives					
Practising					
Emotions					
Shared Positive					
Shared Negative					

*Communcative Functions

5.2b Julie

7 MONTHS (N = 77)

a) Messages attributed by the Mother

At seven months Julie's mother reported one 'assertive' message, two 'receptive' messages and two 'emotional' messages. All the vocalisations classified by the mother as 'Requests', which is an 'assertive message, were 'Requests for Objects'. The 'receptive' messages were either 'self-directed', namely 'Non-Shared Narratives' or 'other-directed', i.e. 'Shared Narratives'. The 'emotional' class included the messages 'Shared Negative Emotion' and 'Shared Positive Emotion'. In the mother's interview appeared also the category 'Squeal'. Although this category does not describe a message, but rather a type of vocalisation it was kept as such in the analysis of the interviews; all the squeals were 'Squeals of Pleasure'.

b) Analysis of Prosody

1) Contour Shape

Vocalisations in all message categories mainly show contours with falling end, and this pattern holds throughout the period of study. As for the overall contour shape, vocalisations interpreted as 'Shared Narrative', 'Shared Positive Emotion' and 'Squeals of Pleasure' are mainly bell shaped, while vocalisations classified in other categories display a variety of contour shapes ($\chi^2 = 7.43$, df = 1, $p<.01$) (Table 5.17).

2) Prosodic Forms

From significant differences in measures of 'peak pitch', 'mean pitch', 'beginning pitch', 'final pitch', 'duration', 'pitch range', 'standard deviation of pitch' and 'rate of pitch change', the following prosodic forms were extracted, each of which expresses a different message of those appearing in this age (Table 5.18).

Form A: Is relatively short, low pitched, non emphatic and has low standard deviation of pitch; also it has the narrowest pitch range of all prosodic forms at seven months. This forms was found in vocalisations interpreted by the mother as 'Requests'.

Form B: Has wider pitch range than Form A and, according to the mother conveys 'Non-Shared Narratives'.

Form C: Is more high pitched than both forms A and B. Form C characterises 'Shared Narratives'.

Form D and E: Are more high pitched than all the previous forms. Moreover, Form E is longer than any other form at seven months. These forms express 'Shared Positive Emotion' and 'Shared Negative Emotion' respectively.

Form F: Characterises the type of vocalisations called *Squeals* and differs from all the other forms in that it is more high pitched, more emphatic, of wider pitch range and high standard deviation of pitch. All the *Squeals* uttered at this age were expressions of pleasure.

3) Conclusion

In conclusion, it can be stated that at seven months Julie's vocal expressions of positive or negative emotions are more high pitched than the utterances conveying any other message. Moreover, the 'receptive', 'other-directed' message 'Shared Narrative' is characterised by a more 'alerting' prosodic pattern than its corresponded 'self-directed' message 'Non-Shared Narrative'.

c) Video Analysis

All the vocalisations recognised by the mother as 'Requests' were 'Requests for Objects' and the analysis of the video records agrees with the mother's interpretation. Similarly, vocalisations classified as 'Non-Shared Narratives' were found in situations where the infant was playing with objects by herself (Table 5.19).

'Shared Narratives' occurred mainly in mother-infant-object situations (83%). On the other hand, vocalisations of any kind expressing 'Shared Positive Emotion' (i.e. 'Squeals' and others) were mostly found in interpersonal situations (60%) (Table 5.19). Moreover, 22 out of 25 (88%) of the vocalisations of 'Shared Positive Emotion' are combined with a smile.

Of the vocalisations expressing 'Shared Negative Emotion' 67% was uttered in situations of potentially negative character, that is in 'Conflict', 'Attention Seeking' or 'Request of Object'. Three vocalisations recognised as conveying 'Shared Negative Emotion' were uttered as Julie was engaged in activity with objects by herself. However, in these cases, in addition to the prosodic form, the mother's interpretation probably relied also on the fact that the infant was alone in the room, as well as on Julie's facial expression of frustration which accompanied only these three vocalisations

out of the 12. When vocalisations of 'Shared Negative Emotion' are combined with the same non-vocal behaviours as vocal 'Requests of Object', 'Shared Narratives' and 'Shared Positive Emotion' vocalisations, the mother recognises negative emotion only on the basis of the prosodic form (Table 5.19).

8 MONTHS (N = 86)

a) Messages attributed by the Mother

At eight months Julie's mother reported for the first time an 'assertive' 'interpersonal' and a 'receptive' 'interpersonal' message, namely 'Attention Seeking' and 'Acknowledgement of Presence' respectively. The remaining messages that appeared at this age, were the same as in seven months; that is, 'Non-Shared Narrative', 'Shared Narrative', 'Request', 'Shared Positive Emotion', 'Shared Negative Emotion', and 'Squeal of Pleasure'.

b) Analysis of Prosody

1) Contour Shape

No message category was correlated with any particular contour shape. For the percentage distribution of message categories to the various contour shapes see Table 5.20.

2) Prosodic Forms

At this age three prosodic forms were extracted on the basis of significant differences in 'peak pitch', 'mean pitch', 'beginning pitch', 'final pitch', 'duration', and 'standard deviation of pitch' (Table 5.21).

Form A: Is relatively long, low pitched and shows low standard deviation of pitch. This form was found in vocalisations interpreted by the mother as 'Shared Narratives', 'Non-Shared Narratives' and 'Requests'.

Form B: This form characterises the 'Squeals of Pleasure'. Vocalisations displaying this form are more high pitched than the vocalisations of the other message categories, except those conveying 'Shared Positive Emotion'. Form B is shorter than both forms A and C described below.

Form C: Differs from Form A in that it is more high pitched, while it is longer than Form B. Form C expresses 'Shared Negative Emotion'.

At eight months the other vocalisations expressing 'Shared Positive Emotion', except the 'Squeals' are shorter utterances than those conveying 'Shared Negative Emotion', and more high pitched than utterances displaying Form A. However, they do

not show significantly different pitch level either from the vocal expressions of 'Shared Negative Emotion' or from the 'Squeals of Pleasure'. Thus, at this age, unlike at seven months, the 'Squeals of Pleasure' were not prosodically distinguished from the other vocal expressions of 'Shared Positive Emotion'.

Vocalisations conveying 'Attention Seeking' are more high pitched compared to those displaying Form A. Moreover, utterances of 'Attention Seeking' have higher standard deviation of pitch than 'Shared Narratives'.

The distribution of vocalisations classified as 'Acknowledgements of Presence' shows great variance regarding all the parameters of pitch height. These vocalisations do not differ on the basis of this feature either from the low pitched Form A or from the high pitched forms B and C. 'Acknowledgements of Presence' are significantly shorter utterances than vocalisations of 'Shared Negative Emotion'.

3) Conclusion

At eight months, 'emotional' and 'interpersonal' messages are expressed by more 'alerting' prosodic forms than 'Shared Narratives', 'Non-Shared Narratives' and 'Requests'. It is worth reminding that at seven months, when 'interpersonal' messages had not appeared, the 'receptive', 'other-directed' message 'Shared Narrative' was characterised by a more 'alerting' prosodic form than its corresponded 'self-directed' message 'Non-Shared Narrative'.

c) Video Analysis

The video analysis demonstrated that in groups of messages which share similar prosodic characteristics, individual messages are distinguished by the mother on the basis of the accompanying non-vocal behaviour and the timing of the vocalisations in relation to other events.

The vast majority of the vocalisations interpreted as 'Shared Narratives' were uttered in situations of 'Converging Interest' (72%). At eight months three kinds of 'Requests' were reported namely, 'Requests for Object', 'Requests for Action on Object' and 'Requests for Action to Self'. As Table 5.22 shows, for most of the 'Requests' the video analysis is in agreement with the mother's interpretation. Two utterances classified as 'Requests for Action on Object' occurred in situations of 'Converging Interest', as do most of the 'Shared Narratives'. In those cases, the cue for the mother's interpretation was probably the timing of the vocalisations, since both of them followed

immediately a change in an 'active' object. 'Non-Shared Narratives' were uttered while Julie was playing with an object by herself without the mother being involved in any way.

The video analysis showed that the 'Attention Seeking' utterances were accompanied by non-vocal behaviour conveying the same message. The majority (66%) of vocalisations recognised as expressing 'Shared Negative Emotion', while similar to the 'Attention Seeking' in prosodic form, occur in different situations that are of potentially negative character, i.e. 'Conflicts' (33%) or 'Conflicts on Object' (33%). Only once was a vocal expression of 'Shared Negative Emotion' accompanied by non-vocal behaviour conveying 'Attention Seeking' (Table 5.22). Only one out of the nine vocalisations conveying 'Shared Negative Emotion' was combined with a facial expression frustration.

Most of the vocalisations interpreted by the mother as 'Acknowledgements of Presence' (63%) were not addressed to a person, but to a person-like object (e.g. a teddy-bear) in situations of 'Converging Interest', 'Joint Interest', or 'Infant-Object' (Table 5.22).

Finally, it seems that Julie's mother relies mainly on situational cues and her facial expression, rather than on the prosodic pattern to differentiate vocal 'Acknowledgements of Presence' from expressions of 'Shared Positive Emotion'. Whenever, 'Acknowledgements of Presence' are uttered in person-person-object situations, they are addressed by Julie to person-like objects, while this is never the case with her vocalisations of 'Shared Positive Emotion'. Moreover, 71% of her vocalisations of 'Shared Positive Emotion' are accompanied by a smile, while this is the case for only 25% of her vocal 'Acknowledgements of Presence'.

9 MONTHS (N = 79)

a) Messages attributed by the Mother

At nine months two new messages appear; specifically, the 'receptive', 'other-directed' message 'Shared Comment', and the 'assertive', 'other-directed', 'referential' message 'Directive'. Vocal 'Directives' manifest Julie's intention to direct the mother's attention to an object of interest, and thus mark her entry into the period of 'Secondary Intersubjectivity'. The other messages that appear at this age are the 'assertive' message 'Request', the 'receptive', 'interpersonal' message 'Acknowledgement of Presence', the 'receptive' message 'Shared Narrative', and the emotional messages 'Shared Positive Emotion' and 'Shared Negative Emotion'; these messages had also appeared at eight months.

b) Analysis of Prosody

1) Contour Shape

Except for the 'Acknowledgements of Presence' which were mainly 'bells', no message category was correlated with any particular pitch contour shape. For the percentage distribution of each message category to the various contour shapes see Table 5.23.

2) Prosodic Patterns

The analysis of prosody revealed the following prosodic forms on the basis of significant differences in measures of 'peak pitch', 'mean pitch', beginning pitch', 'final pitch' (Table 5.24).

Form A: Is relatively short and low pitched. It corresponds to Form A at seven months, and now, in addition to 'Requests', it also conveys 'Shared Comments'.

Form B: Differs from Form A in that it is longer. This pattern characterises vocalisations classified as 'Shared Narratives'.

Form C: Is more high pitched compared to both forms A and B, but is shorter than Form B. Form C conveys vocalisations interpreted by the mother as 'Directives'.

Form D: Is distinguished from the first two forms in that it is more high pitched. Furthermore, Form D is longer than forms A and C. This form expresses 'Shared Positive Emotion' and 'Shared Negative Emotion'.

Vocalisations that the mother recognised as conveying 'Acknowledgement of Presence', like vocalisations expressed by forms C and D, are more high pitched than utterances displaying forms A and B. Moreover, they are shorter utterances than those expressing 'Shared Negative Emotion', but longer than vocal 'Requests'.

3) Conclusion

As at eight months, at nine months 'emotional' and 'interpersonal' messages are expressed by more 'alerting' prosodic forms than 'Shared Narratives', 'Shared Comments' and 'Requests'. However, at nine months the vocal expressions of opposite emotions, that is positive and negative, are not prosodically differentiated. Moreover, now the 'alerting' form conveys the 'referential' message 'Directive', which reflects the infant's developing ability to initiate communication with the mother about an object of interest; this ability characterises the 'Secondary Intersubjectivity' phase.

c) Video Analysis

The video analysis showed that messages conveyed by similar prosodic forms are distinguished on the basis of the accompanying non-vocal behaviour. In particular, vocalisations interpreted by the mother as 'Requests' are combined with gestures conveying the same message, or they occur in situations of 'Conflict with Object', while 'Shared Comments' are uttered in contexts of 'converging' or 'joint' interest. Regarding the 'Directives' and the 'Acknowledgements Presence', the former are mainly combined with 'Directive' gestures, such as 'pointing' or 'showing'. On the other hand, the 'Acknowledgements of Presence' occur mainly in interpersonal situations and they are addressed to the mother, to the stranger or to a person-like object (Table 5.25). At nine months, unlike at eight months, the majority (76%) of 'Acknowledgements of Presence' are addressed to the mother or the stranger, rather than to a person-like object.

The non-vocal behaviours observed in the video record also differentiate the 'Acknowledgements of Presence' from the vocal expressions of 'Shared Positive Emotion'. 'Acknowledgements of Presence' are combined with Julie's responsive behaviours to her mother's invitations for interpersonal play or for play with a person-like object, or with Julie's spontaneous play with a person-like object when she was on her own. On the contrary, her 'Shared Positive Emotion' vocalisations were mainly

accompanied by initiative behaviours by which she invited the mother for interpersonal contact or for a communicative exchange related to an object (Table 5.25). Julie's vocalisations recognised by the mother as expressions of 'Shared Negative Emotion' were found in situations of potentially negative character, i.e. those identified as 'Conflict', 'Conflict with Object' or 'Attention Seeking'. In one case the mother reported that Julie was vocalising 'Shared Negative Emotion', while the accompanying behaviour was 'Unfocused Interest'. However, in this case Julie was alone in the room. In two cases where vocalisations interpreted by the mother as conveying 'Shared Positive Emotion' and 'Shared Negative Emotion' occur in indistinguishable situations. Eight out of twelve vocalisations conveying 'Shared Positive Emotion' are combined with a smile, and 4 out of 14 vocal expressions of 'Shared Negative Emotion' are accompanied by a facial expression of frustration. Moreover, vocalisations of 'Shared Positive Emotion' are not combined with a smile more often than vocalisations conveying any other non emotional message.

Finally, the majority of the 19 vocalisations recognised by the mother as 'Shared Narratives' (74%) were found to conform to the category of 'Converging Interest' in the video analysis (Table 5.25).

10 MONTHS (N = 69)

a) Messages attributed by the Mother

No 'interpersonal' messages such as 'Acknowledgement of Presence' or 'Attention Seeking' are reported at this age. The new messages that appeared at ten months are the 'assertive', 'referential' message 'Shared Interest' and the 'receptive', 'other-directed' message 'Response'. The message 'Response' shows that at the beginning of 'Secondary Intersubjectivity', the mother acknowledges in her infant better awareness of its role in the dialogue between them. At this age were also reported the messages 'Request', 'Directive', 'Shared Comment', 'Shared Narrative', 'Shared Positive Emotion' and 'Shared Negative Emotion'.

b) Analysis of Prosody

1) Contour Shape

Regarding the overall contour shape, the group of messages classified as 'Shared Interest', 'Directives' and 'Responses' (totalling 30) are mainly 'falls' (19), while 'Shared Comments' and vocal expressions of 'Shared Positive Emotion' (totalling 12) are mainly 'bells' (9) ($\chi^2 = 15.85$, df = 1, $p<.001$). 'Shared Narratives' and the vocalisations recognised as conveying 'Shared Negative Emotion' are in the majority 'undulating falls'. The vocal 'Requests' are not related to any particular contour shape (Table 5.26).

2) Prosodic Forms

At ten months, three prosodic forms were extracted on the basis of significant differences in 'peak pitch', 'mean pitch', beginning pitch', 'duration', 'pitch range', and 'standard deviation of pitch' (Table 5.27).

Form A: Is relatively short, low pitched, with narrow pitch range and low standard deviation. This form corresponds to forms A at seven and nine months, but now in addition to 'Requests' and Shared Comments' it also conveys vocalisations interpreted as 'Shared Interest', 'Directives' and 'Responses'.

Form B: Differs from Form A in that it is longer, and with wider pitch range and higher standard deviation. Form B characterises the 'Shared Narratives'.

Form C: Expresses 'Shared Negative Emotion. Is more high pitched than both forms A and B, and also longer and of wider range and higher standard deviation than Form A.

Vocalisations conveying 'Shared Positive Emotion' do not differ in pitch level either from the low pitched forms A and B or from the high pitched Form C. However, they proved to be significantly longer than vocal 'Requests', but shorter than vocal expressions of 'Shared Negative Emotion'. Furthermore, the distributions of pitch range and standard deviation of pitch of the vocalisations classified as 'Shared Interest' show high variance, and they were not statistically differentiated from those displaying forms B and C.

3) Conclusion

At ten months prosody distinguishes three main types of vocalisations: (a) short, 'non-alerting' vocalisations that serve communicative functions and are mainly uttered in situations of 'Joint Interest' or as the infant directs the mother's attention to an object, (b) long, 'non-alerting' vocalisations recognised by the mother as 'Shared Narratives', which were mainly found in situations of 'Converging Interest', and (c) 'alerting' vocal expressions of 'Shared Negative Emotion'.

Moreover, at ten months 3 out of 11 'Shared Narratives' (27%) could be described as 'babbling', that is they were sequences of vocalisations formed by the combination of consonant-and-vowel-like sounds. This percentage will increase further at eleven months.

c) Video Analysis

Vocalisations classified in different message categories, but having the same prosodic form, are distinguished on the basis of the accompanying non-vocal behaviour as well as the timing of the vocalisation. In particular, the video analysis showed that the 'Directives' and the 'Requests' were combined with gestures transmitting or implying the same message. One vocalisation recognised by the mother as 'Request' occurred as the infant was trying to control an active object in a situation of 'Converging Interest'. It is worth noting that in younger ages studied the vocalisations uttered in this situation were expressing 'Shared Negative Emotion'. 'Responses', unlike any other vocalisations, always follow a mother's utterance within an interval of less than one second. The main situational difference between the utterances interpreted as 'Shared Interest' and 'Shared Comments' is that 60% of the former, as compared to only 20% of the latter, occur in situations of 'Joint Interest' (Table 5.28). The report of message 'Shared Interest' at ten months in combination with the fact that these

vocalisations were mainly uttered in situations of 'Joint Interest' constitutes one more evidence for Julie's entry to Secondary Intersubjectivity.

Regarding vocalisations expressing 'Shared Positive Emotion', half of these occur in interpersonal situations, such as 'Responses to Invitations to Play', and the other half were uttered in situations where were found vocalisations with which these share the same prosodic form , namely 'Shared Comments', 'Responses', 'Directives', and vocalisations classified as 'Shared Interest' (Table 5.28). In these cases the mother probably relied for her interpretation on the infant's facial expression, since all the vocalisations conveying 'Shared Positive Emotion' are accompanied by a smile, while this happens only to one 'Response' and one 'Directive'.

Vocalisations of 'Shared Negative Emotion' occur in situations that were observed in the video to have potentially negative character, that is 'Attention Seeking' and 'Conflict'. Finally 'Shared Narratives' occur when there is 'Converging Interest' (Table 5.28).

11 MONTHS (N = 80)

a) Messages attributed by the Mother

At eleven months the message 'Shared Negative Emotion' does not appear. On the other hand, the messages 'Attention Seeking' and 'Non-Shared Narratives', which have not appeared since nine and eight months respectively, are reported again. The remaining messages at this age are the same as at ten months, namely 'Shared Interest', 'Directive', 'Shared Comment', 'Response', 'Shared Narrative', and 'Shared Positive Emotion'.

b) Analysis of Prosody

1) Contour Shape

'Shared Narratives' are mainly of undulating fall shape (67%) the vocalisations in all other message categories display a variety of contour shapes (Table 5.29).

2) Prosodic Forms

Three prosodic forms emerged on the basis of significant differences in 'peak pitch', 'mean pitch', 'beginning pitch', 'final pitch', 'duration', 'pitch range', 'standard deviation of pitch' and 'rate of pitch change (Table 5.30).

Form A: Corresponds to Form A at ten months and conveys the same messages, namely 'Directives', 'Shared Comments', 'Responses' and vocalisations recognised by the mother as expressions of 'Shared Interest'. This pattern is short, low pitched, non-emphatic, with narrow pitch range and low standard deviation of pitch.

Form B: Characterises 'Shared Narratives' and it is longer, more high pitched and of wider pitch range and higher standard deviation than Form A.

Form C: Expresses 'Shared Positive Emotion'. Vocalisations conveyed by this form higher in pitch than the vocalisations of any other message category. Also, it is shorter than Form B, and has wider pitch range and higher standard deviation than Form A.

Vocalisations interpreted by the mother as 'Attention Seeking' are more emphatic than vocalisations in all other message categories, except 'Shared Positive Emotion'. Finally, 'Non-Shared Narratives', like 'Shared Narratives', are long vocalisations compared to those displaying forms A and C, and the vocalisations of 'Attention

Seeking'; however, 'Non-Shared Narratives' are not differentiated in pitch level either from 'Shared Narratives' or from the utterances of Form A.

3) Conclusion

Similarly to ten months, at this age prosody differentiates between short, 'non-alerting' vocalisations that serve communicative functions, long, 'alerting' utterances recognised by the mother as 'Shared Narratives', and vocal expressions of 'Shared Positive Emotion'. The vocalisations in the latter message category are still characterised by a more 'alerting' prosodic form than the vocalisations in the two former message classes.

At this age 58% of the 'Shared Narratives' and 42% of the 'Non-Shared Narratives' can be described as 'babbling' or 'speech-like' sounds.

c) Video Analysis

As at ten months, Julie's vocalisations characterised by Form A, are classified in different message categories, specifically 'Shared Interest', 'Directive', 'Shared Comment' and 'Response'. On the basis of the accompanying non-vocal behaviour and the timing of the vocalisations; Again the difference between Julie's vocalisations interpreted by her mother as 'Shared Interest' and those recognised as 'Shared Comments' is that the former tend to be uttered in 'Joint Interest' situations, while the latter are mainly related to 'Converging Interest' situations (Table 5.31).

The behavioural analysis of 'Attention Seeking' vocalisations is in agreement with the mother's attribution, and 'Non-Shared Narratives' occur while Julie is engaged in an activity with an object without the mother being involved in any way. The 'Shared Narratives' are uttered mainly in situations of 'Converging Interest'. The majority of the vocalisations expressing 'Shared Positive Emotion' were found in 'interpersonal' situations (Table 5.31). Moreover, (91%) of the vocalisations classified by the mother in this category are accompanied by a smile.

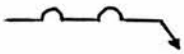
Summary of Julie's Vocalisations

The girl's vocalisations between seven and eleven months showed the following prosodic patterns:

A. Non-Alerting

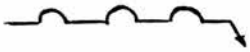
1) Short, low-pitched, monotonous, non-emphatic, with falling contour tail.

2) Medium length, low pitched, fluctuating, non-emphatic, with falling contour tail.



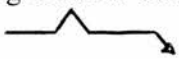
3) Long, low-pitched, monotonous, non-emphatic, with falling contour tail.

4) Long, low-pitched, fluctuating, non-emphatic, with falling contour tail.



B. Alerting

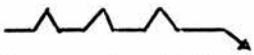
5) Medium length, high-pitched, monotonous, non-emphatic, with falling contour tail.



6) Medium length, high-pitched, fluctuating, non-emphatic, with falling contour tail.

7) Long, high-pitched, monotonous, non-emphatic, with falling contour tail.

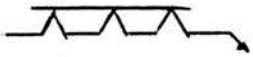
8) Long, high-pitched, fluctuating, non-emphatic, with falling contour tail.



9) Short, very high-pitched, fluctuating, non-emphatic, with falling contour tail.

10) Medium length, very high-pitched, fluctuating, non-emphatic, with falling contour tail.

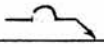
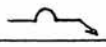
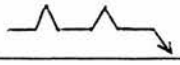
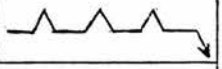
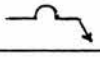
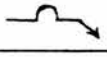
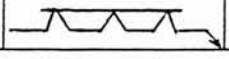
11) Long, very high-pitched, fluctuating, non-emphatic, with falling contour tail.



The distribution of Julie's prosodic patterns by age and class of messages is shown in Table 5.32. As was the case with Robin, before nine months prosodic patterns distinguish between vocalisations occurring in 'mother-infant-object' situations, vocalisations conveying 'interpersonal' messages and those expressing 'emotional' messages. Regarding the first two classes, the 'interpersonal' messages are expressed by more 'alerting' prosodic patterns than the 'referential' messages. After nine months

prosody comes to serve a different role. Specifically, Julie's vocalisations which convey communicative functions and are mainly uttered in 'Joint Interest' and 'Directive' situations, are prosodically distinguished from 'Shared Narratives' which mainly occur in situations of 'Converging Interest' and from vocal expressions of positive or negative emotion.

Table 5.32: Distribution of July's Prosodic Patterns by Age and Class of Messages

MESSAGE	AGE (months)				
	7	8	9	10	11
Assertive					
Interpersonal					
Referential					
Self - Directed					
Other - Directed*					
*Requests**					
Receptive					
Interpersonal					
Referential					
Self - Directed					
Other - Directed					
Shared Narratives					
*Shared Comments**					
*Responses**					
Non-Shared Narratives					
Practising					
Emotions					
Shared Positive					
Shared Negative					

*Communcative Functions

CHAPTER 6

COMPARISONS BETWEEN THE BOY AND THE GIRL

a) Messages attributed by the Mother

Although in most cases Robin's mother and Julie's mother reported the same messages, there are certain messages which differ for the two infants. In particular, Robin uttered vocalisations interpreted by his mother as 'Shared Interest' throughout the period of the study, and vocalisations interpreted as 'Non-Shared Interest' at seven and nine months. On the other hand, Julie uttered 'Shared Interest' vocalisations only after nine months. Her repertoire included vocalisations recognised by her mother as 'Shared Narratives', 'Non-Shared Narratives' and 'Responses'. Robin vocalised 'Shared Narratives' only at nine months, while his mother did not report 'Non-Shared Narratives' and 'Responses' (Table 6.1).

Moreover, Julie combined 'Directive' gestures with vocalisations recognised by the mother as conveying the same message since nine months, while this behaviour was not observed in Robin during the period of the study (Table 6.1). In their study of the development of communicative gestures in two boys Zinober and Martlew (1985) found that joint use of pointing and vocalisation occurred from 14 months.

Table 6.1: Messages reported by Robin's mother (B) and Julie's mother (G) over the period of the study.

MESSAGE	AGE (months)				
	7	8	9	10	11
Non-Shared Interest	**B**		**B**		
Non-Shared Narratives	G	G			G
Attention Seeking	**B**	G	**B**	**B**	**B**G
Invitation to Stranger	**B**	**B**	**B**		**B**
Acknowledgements of Presence		G	G		
Requests	**B**G	**B**G	**B**G	**B**G	
Shared Interest	**B**	**B**	**B**	**B**G	**B**G
Directives			G	G	G
Shared Recognition of Object				**B**	
Responses				G	G
Shared Comments			**B**G	**B**G	G
Shared Narratives	G	G	**B**G	G	G
Shared Positive Emotion	G	**B**G	**B**G	**B**G	**B**
Shared Negative Emotion	**B**G	**B**G	**B**G	**B**G	**B**
Practising					**B**

b) Situations of Occurrence of Infant Vocalisations

After nine months the percentage of 'Joint Interest' and 'Directive' situations where vocalisations are uttered increases dramatically for both infants (Figure 6.1). In other words, at the last quarter of the first year these infants begin to systematically combine non-vocal behaviours expressing joint attention with vocalisations recognised by the mother, in most cases, as conveying communicative functions. However, it is worth noting that this change is manifested at nine months for the girl, but at ten months for the boy. Similarly, Bruner (1983) found that the percent of exchanges initiated by two baby-boys increased dramatically at ten months.

Figure 7.1: Percentage of 'Joint Interest' & 'Directive' Situations by Age

c) Prosodic Patterns and their Functions

Summarising the results presented in detail in Chapter 5, these infants' vocalisations displayed four common prosodic patterns:

PATTERN 1. Short, low-pitched, monotonous, non-emphatic.
PATTERN 2. Long, low pitched, fluctuating, non-emphatic, with falling contour tail.
PATTERN 3. Long, high pitched, non-emphatic.
PATTERN 4. Medium, high pitched, non-emphatic.

During the period of the study, i.e. between 30 and 50 weeks the following changes were observed regarding the form and the function of these patterns:

PATTERN 1: Its form remained the same; however, its function changed after nine months to serve the expression of the newly developing motives of Secondary Intersubjectivity.
PATTERN 2: This pattern conveys the same message throughout the period of the study, but its form becomes progressively more speech-like.
PATTERNS 3 and 4: Neither their form nor their function change over the period of the study.

Pattern 1 appears in boy's vocalisations at seven and eight months conveying 'Shared Interest' in situations of 'Converging Interest', it disappears at nine and ten months and then reappears at eleven months expressing 'Shared Interest' but now in situations of 'Joint Interest'. The same pattern is found in girl's vocalisations at seven months conveying 'Requests' and again at nine, ten and eleven months conveying 'Directives' as well as 'Requests', 'Shared Interest', 'Shared Comments' and 'Responses' which occur mostly in situations of 'Joint Interest'.

Pattern 2 was found in girl's vocalisations at seven, eight, ten and eleven months. At all ages it was conveying 'Shared Narratives' in situations of 'Converging Interest'. However, this pattern's duration increased with age ($F(4,91)=20.59$, $p<.001$). Moreover, at ten months 4% of the vocalisations displaying this pattern were babbling; this percentage increased to 15% at eleven months. The boy uttered vocalisations showing Pattern 2 at nine and eleven months. In the first case, these vocalisations were judged by the mother as 'Shared Narratives' and occurred mainly in situations of 'Converging Interest'. At eleven months the vocalisations displaying this prosodic pattern were judged as 'Practising' and 17% of them were babbling.

For the new form of Pattern 2, which appeared after ten months both mothers often commented that their infant "imitated speech". Moreover, anecdotal observations showed that in many cases the mothers matched and expanded these utterances in the direction of the adult form. Infant vocal imitation and maternal contingent response are considered as important mechanisms of vocal learning (Locke, 1995). Relevant studies present discrepant results on the infants' capacity to imitate the pitch contour of their speaking partners (Siegel et al, 1990; Masataka, 1992). Dodd (1972) demonstrated that nine-to-twelve month old infants were more likely to produce vocalisations that included consonants after listening to their own live babble than live speech. These findings and the observations of the present study suggest that vocal learning involves more complex processes than vocal imitation as infants progress down the path to spoken language. It seems that mothers selectively reinforce infant utterances which belong to infant's spontaneous repertoire by matching, expanding or transforming them in the direction of the target language (Locke, 1995).

Patterns 3 and 4 that expressed emotions of any kind, which regulate communication did not show any change over the period of the study.

In summary, Pattern 1 which remains unchangeable between 30 and 50 weeks, functions in situations of 'Converging Interest' and after nine months in situations of 'Joint Interest'. On the other hand, the new form of Pattern 2, which appears after ten months is uttered in situations of 'Converging Interest', as is the case with its antecedent form. Thus, it seems that new forms of prosodic expression occur in psychologically mature situations, while the psychologically novel situations are expressed by mature prosodic forms.

d) Maturation of the Vocal System

1) Prosodic Features

A linearity test was applied to detect any developmental changes in the prosodic features that are considered to be indices of maturation of the vocal tract independent of message. These prosodic features are 'peak pitch', 'mean pitch', 'pitch range' and 'standard deviation of pitch' (Sheppard and Lane, 1968; Laufer and Horii, 1977; Keating and Buhr, 1978; Robb and Saxman, 1985; Oller, 1980; Stark, 1980).

The results showed for both infants a general tendency of decrease in 'peak pitch' [(F (4,318) = 6.62 p<.01-boy) (F(4,387) = 44.68 p<.001-girl)] and mean pitch [(F(4,318) = 7 p<.001-boy) (F(4,387) = 43.83 p<.001-girl)] for both infants during

the period of study, from 30 to 50 weeks. Moreover, for the vocalisations of the girl exhibited a decrease in 'pitch range' ($F(4,387) = 45.03$ $p<.001$) and in 'standard deviation of pitch' ($F(4,387) = 31.59$ $p<.001$), whilst for the boy the values of these features showed no significant change.

Robb and Saxman (1985), who obtained fundamental frequency values from the spectrogram for fourteen children between 11 and 25 months, also demonstrated a decrease in average fundamental frequency and in variability of pitch as a function of age. The results were explained by the physical maturation of the infants observed at this period.

2) Babbling

The girl started systematically uttering babbling vocalisations at ten months and these constituted 4% of the total vocalisations analysed at this age. At eleven months the percentage of babbling vocalisations increased to 15%. On the other hand, in the boy babbling vocalisations were first observed at eleven months; these were 17% of the total vocalisations analysed at this age. This finding may be an index of earlier vocal maturation of the girl.

Correlations between Prosodic Variables

Correlations between the measures of peak pitch, mean pitch, beginning pitch, final pitch, duration, pitch range, standard deviation of pitch, and contour shape were calculated and the results are shown in Table 6.3[1] . Beginning pitch was significantly positively correlated with both peak pitch and mean pitch. Robb and Saxman (1985) found significant correlation between the variables of mean fundamental frequency and fundamental frequency onset ($r = 0.98$ $p<.01$). The authors make the remark that the high positive correlation supported previous observations that measures of average fundamental frequency and fundamental frequency onset yield essentially the conclusions about pitch level.

Furthermore, it was found that contour shape was positively correlated with duration, pitch range and standard deviation of pitch. Specifically, contour shapes which emerge from more than one perceivable changes in pitch, namely 'bell', 'undulating fall' and 'undulating rise' appear in long prosodic patterns, with wide pitch range and high standard deviation of pitch. Delack and Fowlow (1978) note that although contours are supposed to be construed as complex composites of more basic features, they are not uniquely determined by these; in fact, the converse could be true. In other words, the observed differences in contour distributions cannot be attributed solely to concomitant fluctuations in variables such as duration, pitch range and standard deviation of pitch. Moreover, whether the infant is manipulating contours or their associated features, has not been systematically examined from either psychological or physiological point of view (Delack and Fowlow, 1978). A meaningful interpretation of the above findings awaits further specification.

Duration of the utterance was not correlated with pitch range or standard deviation of pitch, as was expected. This result is in agreement with Robb and Saxman (1985), who found no correlation between MLU (mean length of the utterance) and the inter-utterance standard deviation of fundamental frequency,

---

[1] Correlations between the continuous variables (i.e. peak pitch, mean pitch, beginning pitch, final pitch, duration, pitch range and standard deviation of pitch) were calculated using the Pearson's product moment correlation coefficient (r). In order to calculate the correlation between a continuous variable and a nominal variable (in this case contour shape) the latter has to be dichotomous. Thus, the categories used for describing contour shape were grouped into two classes: Class A included contour shapes which emerge from more than one perceivable changes in pitch, namely 'bell', 'undulating fall' and undulating rise'. Class B included contour shapes which emerge from none or just one perceivable change in pitch, namely 'level', 'rise' and 'fall'. The dichotomous variable was then correlated with the continuous variables using the Biserial correlation coefficient (r_b) (Howell, 1992; Coolican, 1994).

although the correlation coefficient between segment duration and inter-utterance standard deviation of fundamental frequency approached criterion significance (r = 0.62 exceeding the .05 level).

Table 6.3: Intercorrelation matrix of eight variables for the two infants.

	Peak Pitch	Mean Pitch	Beginning Pitch	Final Pitch	Duration	Pitch Range	Standard Deviation of Pitch
Peak Pitch	-----						
Mean Pitch	**0.834***	-----					
Beginning Pitch	**0.77***	**0.735***	-----				
Final Pitch	0.051	0.014	-0.002	-----			
Duration	-0.015	-0.02	-0.023	-0.011	-----		
Pitch Range	-0.069	-0.018	-0.017	-0.277	0.166	-----	
Standard Deviation of Pitch	-0.081	-0.005	-0.045	-0.268	0.154	**0.933***	-----
Contour Shape	-0.025	-0.063	-0.013	-0.04	**0.907***	**0.862***	**0.751***

* $p<.01$

CHAPTER 7

DISCUSSION AND CONCLUSIONS

The present study was carried out with two purposes. First, to obtain a descriptive or ethological account of the prosodic characteristics in vocalisations of infants between 30 and 50 weeks of age, and second, to examine the role of different prosodic forms in conveying meanings in combination with other (visible) expressive modalities.

The results concerning the role of infant prosody in the second half of the first year can be stated as follows:

(i) Specific prosodic patterns do not relate one to one with particular messages; rather they distinguish broad functional classes of messages, namely 'referential', 'interpersonal', 'assertive', 'receptive' and 'emotional'.

(ii) Infant prosody does not primarily convey identifiable linguistic functions, but in combination with the other expressive modalities, it reflects the infant's developing motivational organisation.

(iii) After nine months infant prosody contributes to the expression of motives for joint attention which are becoming active this period; joint attention is considered to be one of the prerequisites for the acquisition of adult language.

These results will now be discussed in turn.

(i) Individual messages within each of the three classes referred to above (i.e. 'referential', 'interpersonal', 'assertive', 'receptive' and 'emotional') were distinguished by the mothers on the basis of the accompanying expressive non-vocal behaviour, the timing of a vocalisation in relation to changes in the mother's activity at the time or in 'active' object that was the focus of the infant's interest, or their knowledge of the infant's previous experiences in play routines. This set of findings constitutes further evidence on complexity in the network of communication between the mother and the infant, wherein meanings are generated and shared. Anyone studying the meaning of infant vocalisations will need to take this complexity into account when attempting to interpret their communicative functions.

This finding is in agreement with studies on adult intonation, which support the argument that a particular intonation contour is not related to a particular meaning, but rather it can gain different meanings depending on the context in which it is uttered (Cuttler, 1977; Pierrhumbert and Hirschberg, 1990). Traditionally, students of adult language have focused on the phonological use of pitch movement, that is on 'intonation', and they have attempted to establish one-to-one relations between certain intonational contours and particular communicative attitudes of the speaker, such as 'politeness', 'incredulity' or 'irony' (Lieberman, 1975), or particular speech acts, such as 'statements', 'requests' or 'contradictions' (Lieberman and Sag, 1974; Couper-Kuhlen, 1986). However, the many-to-one mapping between attitudes or speech acts and intonation contours suggests that the expressed meaning of an utterance is better understood as derived from the contour meaning interpreted in a broader behavioural context (Pierrhumbert and Hirschberg, 1990).

In the same vein, Ladd and his colleagues (1985) examined experimentally the function of 'intonation contour', 'voice quality', and 'fundamental frequency' range in signalling affect. According to the authors, there may be two different types of vocal cues which contribute to the signalling of a speaker's states and intentions in essentially different ways. These cues are (a) the continuous acoustic variables whose variation is more or less directly correlated with variations in the affective message, and (b) those that are organised into linguistic (and perhaps perceptual) categories whose interpretation depends on interaction with other cues in the context, including the verbal content (e.g., a falling intonation contour is judged as neutral with a *wh-* question, but aggressive or challenging with a *yes/no* question). The linguistic variables tend to signal the speaker's attitudes with a greater cognitive or attitudinal component than the messages conveyed by the continuous prosodic variables.

In Ladd's study (Ladd et al, 1985) listeners judged three sentences with not obviously 'loaded' verbal content on the basis of two separate rating forms, one for arousal-related states and the other for more cognitive attitudes. The major result of the experiment is that 'fundamental frequency range' and 'voice quality' had a strong effect on judges' inference of speaker 'arousal'. Yet there are also significant effects of voice quality and fundamental frequency range on more 'cognitive' attitudes (e.g. 'emphasis', 'contradiction' and 'reproach'), although these effects are all weaker than those that signal arousal states. On the other hand, intonation contour does have a significant effect on the cognitive-attitude scales: an uptrend signals greater emphasis, stronger contradiction, and less cooperativeness. However, intonation

contour also has effects on the arousal scales that are sometimes stronger than those on the cognitive-attitude scales. Uptrending intonation, like wide fundamental frequency range and harsh voice quality, is interpreted as signalling 'arousal', 'annoyance' and 'involvement'. Thus, the distinction between arousal, or level of motivation, and cognitive-attitudes is not directly reflected in the acoustic cues.

The authors conclude that the results are in agreement with the assumption, now widespread in the literature on linguistic pragmatics, that the interpretation of an utterance is the result of an active process of inference based on all the information - verbal, paralinguistic, and contextual - available to the listener (Ladd et al, 1985). It seems that prosodic features in adult language are not primarily linguistic entities serving linguistic functions; rather, they constitute an essential part of a coherent system involving also other non-linguistic expressive behaviours such as gaze and gestures; this expressive system conveys states of arousal, mental states and communicative functions. The findings of the present study, on the role of prelinguistic prosodic patterns in mother-infant communication in the second half of the first year, show that this whole expressive system is functional in humans from very early in life, before language is acquired.

Infants express themselves with coordinated behaviours of their whole body from the first 'Protoconversations' (Bateson, 1975; Trevarthen, 1982). Halliday (1975, 1978) named the infant expressive system as it appears in the last quarter of the first year 'Protolanguage'. 'Protolanguage' was described as combinations of vocalisations, gestures, body postures and gaze, to signal wishes, feelings, purposes and experiences in relation to those of the partner. Most children use combinations of all the expressive modalities, although they may show a preference for a particular one, i.e. prosody or gestures. 'Protolanguage' is characterised by great versatility and is influenced by imitation of how other people express themselves, and recognition of what they attend to. It is not speech in a recognised mother tongue; however, certain infant' vocalisations (or gestures) may become repeated conventions in the family (Trevarthen, 1990). Halliday (1978) emphasises that the essential ingredient of 'Protolanguage' is not the form of the output, but rather the nature of the 'Act of Meaning' it expresses. An 'Act of Meaning' is a communicative act that is intentional and symbolic. A symbolic act is one of which the meaning and success criteria do not reside on its own performance; it requires interpretation or understanding by the communicative partner (Halliday, 1978).

Research on animal signals, following Darwin, shows that animals possess a similar complex expressive system. Marler and his colleagues (1992) investigated the relationships between acoustic characteristics of signals, properties of the classes of stimuli with which signals are associated, and motivational state at the time of signal production, in an attempt to provide more insightful comparisons between the communication systems of non-human and human animals. The authors distinguished two kinds of signals namely, the 'motivational' and the 'referential'. Signals traditionally judged as 'referential' encode and transmit information about external stimuli that are highly specific. On the contrary, 'motivational' signals convey information about the motivational state of the animal. No animal signal that has been identified as referential is devoid of motivational content, and the opposite is also true.

It has been suggested that call morphology is determined by either motivational state or stimulus characteristics. However, given the fact that most natural vocal signals have both referential and motivational characteristics- even though the proportional contributions of these factors vary widely- it is claimed that most probably the acoustic morphology of calls is determined by a complex interaction between stimulus attributes and motivation. Thus, a vocal signal cannot be interpreted on the basis of the signal characteristics alone; in decoding a signal receivers take into account contextual cues provided by the environment and/or by the non-vocal behaviour of the sender (Marler et al, 1992).

The fact that prosody in infant vocalisations of the second half of the first year, adult utterances, as well as in vocalisations of non-human primates conveys primarily emotions, mental states and communicative functions in cohesion with other non-vocal expressive behaviours, compels one to study not only to what extent infant prosody is similar to or different from to adult intonation - as has been happening so far -, but also to what extent the infant expressive system is idiosyncratic, language dependent, universal, or even species-specific.

(ii) The findings about the role of infant prosody can be interpreted by the Theory of Intersubjectivity which explains psychological development as a result of changes in innate motives for communication. Motives are described as the core of the subject's intrinsic mental organisation, or otherwise as interior processes of the brain which anticipate and interpret consequences of behavioural action, without, however, necessarily taking into account the limitations or advantages of the actual external reality for their expression. Motives are distinct from emotions which transmit the

colour of motives to the others, so they can know what the subject wants (Trevarthen, 1982). According to Trevarthen (1982), psychological development results from changes in motives themselves, as well as from changes in the power of perceptual and motoric mechanisms which serve the expression of motives. Moreover, the development of motives depends on experience and in intimate communication with the caregiver. Trevarthen (1982) allows for two fundamental motives: the motive for interpersonal exchange with other individuals and the motive for gaining knowledge about objects. In the course of development the fundamental motives are combined to form new more complex ones.

Before nine months, motives for person-person exchanges ('interpersonal') and motives for exploration of the surrounding inanimate environment are functioning separately ('referential'). In other words, infants are not able to follow or direct the other's focus of attention. The present study demonstrated that at seven and eight months prosody distinguishes between vocalisations recognised by the mothers as conveying 'interpersonal' messages, 'assertive' or 'receptive', from vocalisations which are attributed 'referential' messages or occur in situations of 'Converging Interest'.

At nine months a reorganisation takes place that will lead to the functional combination of motives for gaining knowledge about objects with motives for interpersonal communication. A development of 'Self-Other' awareness was observed which lays the groundwork for joint attention. Infants overtly gain a more sensitive insight into the 'Other's' conscious and become capable of a more efficient control of intersubjective orientation which is the core of joint attention; they convey a clear quality of interests, intentions and emotions in relation to the interest, intentions and feelings of the other. Infants of this age become markedly more interested in the purposes of others as indicated in their actions, expressive movements and utterances, and are able to follow their focus of attention. Moreover, at this age infants take over the role of agent and initiate themselves communicative exchanges with the mother about objects of interest.

Julie's vocalisations accompanied by non-vocal behaviours which direct the mother's attention towards an object, or manifest 'Joint Interest', that is voluntary adjustment to the mother's interests and intentions, display different prosodic patterns from vocalisations combined with non-vocal behaviours which express 'Converging Interest', i.e. passive acceptance of the mother's participation in her activity with an object.

On the other hand, Robin's vocalisations after nine months found to distinguish between 'assertive' messages by which he initiates a communicative exchange with the mother, and 'receptive' messages which function as a response to the mother's communicative activity. The 'assertive' messages displayed more 'alerting' prosodic patterns than the 'receptive' ones. The distinction observed between 'assertive' and 'receptive' messages conveyed by means of prosody reveals a development in 'Self-Other' awareness.

One can conclude that over the period studied infant prosody proved to be primarily a means of manifestation of the innate motives for communication, which are considered to be the core of psychological development. Especially after nine months prosody communicates changes in awareness or expresses purposes in intersubjective exchanges with the mother. The objective of studies on infant prosody should not be restricted to detection of similarities and differences with the intonation system of the target language. Rather, infant prosodic repertoire needs also to be studied as an autonomous expressive means which plays a crucial role in the regulation of communication with the caregiver.

Previous studies of infant prosodic patterns in the second half of the first year have also showed that messages which demand some intervention by the partner differ prosodically from messages which do not demand such an intervention, the first being marked by rising melodic patterns and high pitch (Halliday, 1975; Furrow, 1984; Marcos, 1987; D'Odorico et al, 1991). These authors simply describe this expressive behaviour, without, however, attempting to explain its manifestation in terms of the psychological processes which generate it.

The finding that prosodic patterns in infant vocalisations reflect the motivational development of an infant is complementary to the results of studies on prosody in maternal speech to infants. In particular, it has been shown that, as the infant grows older, mother's prosodic patterns change so as to convey new communicative intentions which match with the infant's communicative intentions that are dominant at each developmental stage (Marwick et al, 1984; Trevarthen and Marwick, 1986). In the last quarter of the first year, when the infant is capable of cooperative play, large transformations occur in the mother's behaviour as 'utterer'. Her spoken messages change and 'questions' and 'invitations' are replaced with 'imperatives' and 'directives'. Except the verbal content, this change is also conveyed

by modifications in the prosodic patterns of her vocalisations (Trevarthen and Marwick, 1986).

(iii) Just before the emergence of the first words, certain prosodic patterns in combination with other modalities were shown to serve the expression of joint attention. Development of joint attention is crucial for the emergence of language.

Messer (1997) stated that recurrent episodes of joint attention help the infant to determine adults' attentional focus and, thus, the intended referent of their language. This author attributes any psychological changes to learning from recurrent interactive episodes. However, studies on the development of cooperative action (Trevarthen and Hubley, 1978; Hubley and Trevarthen, 1979; Hubley, 1983) as well as the finding of the present study that infant prosody contributes to the expression of cooperative awareness, suggest that infants actively seek themselves joint attention after nine months.

Tomasello and Farrar (1986) explored systematically the role of joint attentional processes in language acquisition in 24 children at 15 and 21 months. The results showed that inside, as opposed to outside, joint attentional episodes children produced more utterances, more words and more words referring to objects per minute. Differences between the values inside and outside joint attentional episodes were greater at 21 months. Like their children, mothers produced more utterances per minute inside as opposed to outside joint attentional episodes. Further, inside 'joint attention' episodes, maternal references to objects that were already the child's focus of attention were positively correlated with the child's vocabulary at 21 months.

The expression of joint attention by means of prosody and other expressive modalities facilitates the development of the more complex motives for symbolic play in the second year. Symbolic play supports the learning of cultural conventions and maternal language. After the first year, infants become capable of selective retention and imitation of experiences about the ways in which significant others use the world. The memories of these experiences are retained because interest in them has been shared in episodes of joint attention. Gradually, the internalised imagination of shared roles, acts and tasks gains a spontaneous self-creating character, by means of which a child may make dynamic meanings while playing alone. This is the entry to the world of 'fantasy' play. It is clear form the imaginative play of toddlers with their mothers that motivation for creating meanings and for

coding objects and experiences in conventional categories, or symbols are part of the child's thinking before words are mastered. The cooperative learning of language needs flexibility of imagination and purposes, which is expressed in pretend play. Thus, around the middle of the second year infants have the cognitive abilities and the motivation to take up conventional use of tools, roles and rituals of performance, in order to retain and cultivate the companionship with the others, and become active members of their cultural community. In turn, social cooperation is enriched by language through reference to things of a different time and place, and elaboration of detailed imaginative narratives and 'dramas'.

Furthermore, Bruner (1975, 1978) and Halliday (1973) have argued that joint attention may be important for the acquisition of grammar. Infants are endowed with the ability to actively participate in episodes of joint attention. Their expressive repertoire, including prosody allows them to convey functions such as 'predicting the environment', 'interacting transactionally' or 'getting to goals with the aid of another person'. These functions are related to innate and universal concepts of action, namely 'agent', 'object', 'instrument', 'possession' or 'location'. Such concepts correspond to grammatical categories of the adult language. Grammatical forms might then be the surface structure of language, depending for their acquisition on a prior understanding of deep proto-semantic concepts about action. This understanding is facilitated by joint attention. Nevertheless, Bruner himself admits that "the issue of whether rules of *grammar* can somehow be inferred or generalised form the structure of our knowledge of the world is a very dark one" (Bruner, 1983 p. 34).

Language is not generated by a Language Acquisition Device innate in human beings, which output is a set of syntactic rules. Rather, it requires a sensitivity to a sound system, to structural constraints, to referential requirements, and to communication objectives. These developments are cultivated in communicative exchanges with the caregiver which are first fulfilled by prelinguistic means of expression such as prosody (Bruner, 1978).

Previous relevant studies because of their design could not have revealed the role of prosody that was observed in the present research. Unlike this study, where infant vocalisations were attributed a message according to what the mother felt her infant was expressing in each case, in the studies of D'Odorico and Franco (1991) and Halliday (1975) infant vocalisations were interpreted on the basis of predetermined categories which did not include emotions. Moreover, in Furrows

(1984) research prosodic features in infant vocalisations were measured only in interpersonal situations, while situations where the mother and the infant are engaged in an activity with an object of shared interest were excluded. Nevertheless, these routines are the typical ones of this period and occupy the major part of the infants' active time.

This research demonstrated that both infants, the boy and the girl entered the Secondary Intersubjectivity phase at nine months, although this development was manifested in different ways in the two cases. Further research on a large number of infants needs to address the question of gender differences in the development of cooperative awareness and in the role of prosody in relation to the other modalities in its expression. However, this two-case study suggests that girls may be ahead of boys in vocal maturation, since the girl showed earlier and more often babbling vocalisations. Fenson and his colleagues (1993), who claim continuity between babbling and the first words, found a tendency for girls to acquire words at a slightly faster rate than boys. Moreover, it has been shown that boys tend to be developing a few weeks later than girls around three to six months in auditory evoked responses (Schucard et al, 1984). This finding has been explained by an earlier passage through a transitory change of asymmetry in the females than in the males (Trevarthen, 1996).

In conclusion, the findings of the present study support the idea that human beings are endowed with an expressive prosodic repertoire that is functional from the beginning of life, and which, during the second half of the first year, reflects primarily the infants' developing motivational organisation and after nine months contributes to the expression of joint attention. Joint attention constitutes one major prerequisite for the acquisition of language. Part of the infants' prosodic repertoire, which is a compound of idiosyncratic and imitative forms, will develop into the intonation system of the target language, whilst another part continues to be the means for the expression of unconventional, and possibly universal, human mental states, communicative functions and emotions.

The idea I tried to support in the present thesis through scientific findings regarding the vocal expressions of two young infants, is brilliantly described by the Greek author Nikos Kazantzakis:

"[Zorbas] this primitive man, was going beyond the poor crust of what we call civilisation....reaching the essence of life.....When he really wanted to express

himself, his throat was returning to the ages before the existence of human beings, when an unarticulated sound was an unrivalled composition which comprised whatever nowadays we call music, poetry and passion".

Nikos Kazantzakis 'Alexis Zorbas'

REFERENCES

Ambercombie, L. (1923) Principles of English Prosody. London: Secker.

Austin, J.L. (1962) How to Do Things with Words. Oxford: Clarendon Press.

Bakeman, R. and Adamson, L.B. Coordinating attention to people and objects in mother-infant and peer-infant interaction. Child Development, 55, 1278-1289.

Barry, W.J. (1981) Prosodic functions revised again. Phonetica, 38, 320-340.

Bates, E., Camaioni, L. and Voltera V. (1975) The acquisition of performatives prior to speech. Merill Palmer Quarterly, 21, 205-226.

Bates, E. (1976) Language and Context: Studies in the acquisition of pragmatics. New York: Academic Press.

Bates, E., Benigni, L., Bretherton, I., Camaioni, L. and Voltera, V. (1977) From gesture to the first word: On cognitive and social prerequisites. In M. Lewis and L. Rosenblum (eds.), Interaction, Conversation and the Development of Language. New York: Wiley.

Bateson, M.C. (1975) Mother-infant exchanges: The epigenesis of conversational interaction. In D. Aaronson and R.W. Rieber (eds.), Developmental Psycholinguistics and Communication Disorders. New York: New York Academy of Sciences.

Bendor-Samuel, J.T. (1960) Some problems of segmentation in the phonological analysis of Terena. In F.R. Palmer (ed.), Prosodic Analysis, 214-21. London: Oxford University Press.

Bernstein, L (1976) The unanswered question. The Charles Eliot Norton Lectures. Cambridge, Mass: Harvard University Press.

Best, C.T., Hampel, L. and Chiang, S.Y. (1992) Discrimination of English and Mandarin Discourse Prosody Contrasts by English Learning Infants. Poster

presented at Biennial Meeting of the International Conference on Infant Studies, Miami FL.

Blasdell, R. and Jensen, P. (1970) Stress and word position as determiantns of imitation in first language lerners. Journal of Speech and Hearing Research, 13, 193-202.

Bloch, B. and Trager, G.L. (1948) Outline of Linguistic Analysis. Baltimore: Special Publication of the Linguistic Society of America.

Bloom, L. (1973) One Word at a Time: The use of single word utterances before syntax. The Hague: Mouton.

Bolinger, D. (1964) Around the edge of language: Intonation. In D. Bolinger (ed.), Intonation, 19-29. Harmondsworth: Penguin.

Bolinger, D. (1989) Intonation and its Uses: Melody in grammar and discourse. London: Edward Arnold.

Brown, C.J. (1979) Reaction of infant's to their parents voices. Infant Behaviour and Development, 2, 295-300.

Bruner, J. (1975) From communication to language- A psychological perspective, Cognition, 3, 255-287.

Bruner, J. (1978) On prelinguistic prerequisites of speech. In R. Campbell and P. Smith (eds.), Recent Advances in the Psychology of Language, 199-251. New York: Plenum.

Bruner, J. (1982) The beginnings of request. In K.E. Nelson (ed.), Children's Language, vol. 3, 91-137. Hillsdale, N.J. Erlbaum.

Bruner, J. (1983) Child's Talk: Learning to use language. Oxford: Oxford University Press.

Burnet, J. (Lord Monboddo) (1773-1792) The Origin and Progress of Language. Edingurgh: Creech.

Campbell, M. and Greated, C. (1987) The Musician's Guide to Acoustics. J.M. Dent & Sons Ltd: London, Melbourne.

Carpenter, R.L., Mastergeorge, A.M. and Coggins T.E. (1983) The acquistition of communicative intentions in infants eight to fifteeen months of age. Language and Speech, 26, 101-116.

Carrell, J. and Tiffany, W.R. (1960) Phonetics: Theory and application to speech improvement. New York: McGraw-Hill.

Carter, A.L. (1978) The development of systematic vocalisations prior to words. A case study. In N. Waterson and C. Snow (eds.), The Development of Communication, 127-138. New York: Wiley.

Chiang, S. (1992) Chinese Infant Perception of Native and Non-Native Prosodic Contrasts. Poster presented at Biennial Meeting of the International Conference on Infant Studies, Miami, FL.

Chomsky, N. (1959) Review of B.F. Skinner's verbal behaviour. Language, 35, 16-58.

Chuang,C.-K and Wang, W.S.-Y (1978) Psychological pitch biases related to vowel quality, intensity difference, and sequencial order. Journal of the Acoustical Society of America, 64, 1004-1014.

Condon, W.S. and Sander, L.S. (1974) Neonate movement is synchronised with adult speech: Interactional participation and language acquisition. Science, 183, 99-101

Coolican, H. (1994) Research Methods and Statistics in Psychology. Hodder and Stoughton

Couper-Kuhlen, E. (1986) An Introduction to English Prosody. London: Edward Arnold

Cruttenden, A. (1986) Intonation. Cambridge: Cambridge Universtity Press.

Crystal, D. (1969) Prosodic Systems and Intonation in English. Cambridge: Cambridge University Press.

Crystal, D. (1973a) Non-segmental phonology in language acquisition: A review of the issues. Lingua, 32, 1-45.

Crystal, D. (1973b) Linguistic mythology in the first year of life. British Journal of Disorders of Communication, 11, 29-36.

Crystal, D. (1975) The English Tone of Voice. London: Edward Arnold.

Crystal, D. (1979) Prosodic development. In P. Fletcher and M. Garman, Language Acquisition, 33-48. Cambridge: Cambridge University Press.

Crystal, D. and Quirk, R. (1964) Systems of Prosodic and Paralinguistic Features in English. The Hague: Mouton.

Cuttler, A. (1977) The context dependence of 'intonational meaning'. Papers from the 13th Regional Meeting of the Chicago Linguistic Society, 104-115.

Dodd, B. (1972) Effects of social and vocal stimulation on infant babbling. Developmental Psychology, 11, 478-484.

D'Odorico, L. and Franco, F. (1991) Selective production of vocalisation types in different communication contexts. Journal of Child Language, 18, 475-499.

Darwin, C. (1872) The Expression of the Emotions in Man and Animals. London: John Murray.

Darwin, C. (1977) A biographical sketch of an infant. Mind, 2, 285-294.

de Laguna, G. (1927) Speech: Its function and development. New Haven: Yale University Press.

De Angulo, J. (1929) Grammatical processes: incremental vs. autonomic. Language, 5, 117-118.

de Boysson-Bardies, B., Bacri, N., Sagart, L. and Poizat M. (1982) Timing in late babbling. Journal of Child Language, 8, 525-539.

de Boysson-Bardies, B., Sagart, L. and Durand, C. (1984) Discernible differences in the babbling of infants according to target language. Journal of Child Language, 11, 1-15.

DeCasper, A. and Fifer, W.P. (1980) On human bonding: newborns prefer their mothers' voices. Science, 208, 1174-1176.

DeCasper, A. and Spence, M. (1986) Prenatal maternal speech influences mewborns' perception of speech sounds. Infant Behaviour and Development, 9, 133-150.

Delack, J.B. and Fowlow, P.J. (1978) The ontogenesis of differential vocalisation: development of prosodic contrastivity during the first year of life. In N. Waterson and C. Snow (eds.), The Development of Communication, 93-110. New York: Wiley.

Dore, J. (1975) Holophrases, speech acts and language universals. Journal of Child Language, 2, 21-39.

Dore, J., Franklin, M.B., Miller, R.T. and Ramer, A.L. (1976) Transitional phenomena in early language acquisition. Journal of Child Language, 3, 13-28.

Dore, J. (1983) Feeling, form and intention in the baby's transition to language. In R. Golinkoff (ed.), The Transition from Prelinguistic Communication, 234-253. Hillsdale, N.J. Earlbaum.

Eady, S.J. (1982) Differences in the Fo patterns of speech: Tone language versus stress language. Language and Speech, 25, 29-42.

Eimas, P.D. (1974) Auditory and linguistic processing for place of articulation by infants. Perception and Psychophysics, 16, 513-521.

Ferguson, C.A. (1964) Baby talk in six languages. Americal Anthropologist, 66, 103-114.

Fernald, A., Taeschner, T., Dunn, J., Papousek, M., de Boysson-Bardies, B. and Fukui I. (1989) A cross-language study of prosodic modifications in mothers' and fathers' speech to preverbal infants. Journal of Child Language, 16, 477-501.

Fernald, A. and Simon, T. (1984) Expanded intoantion contours in mothers' speech to newborns. Developmental Psychology, 20, 104-113.

Ferrier, L.J. (1985) Intonation in discourse: Talk between 12-month-olds and their mothers. In K.E. Nelson (ed.), Children's Language, vol. 5, 35-60. Hillsdale, N.J. Erlbaum.

Field, T.M. (1985) Neonatal perception of people: Maturational and individual differences. In T.M. Field and N. Fox (eds.), Social Perception in Infants. Norwood N.J.: Ablex.

Fitchen, M. (1931) Speech and music development of a one year old child. Child Development, 2, 324-326.

Flax, J., Lahey, M., Harris, K. and Bothhroyd, A. (1991) Relations between prosodic variables and communicative functions. Journal of Child Language, 18, 3-19.

Fónagy, I. and Magdics, K. (1963) Emotional patterns in intonation and music. In D. Bolinger (ed.), Intonation, 19-29. Harmondsworth: Penguin.

Fraisse, P. (1982) Rhythm and tempo. In D. Deutch (ed.), The Psychology of Music. New York: Academic Press.

Fridman, R. (1980) Proto-rhythms: Nonverbal behaviour in language and musical acquisition. In M.R. Key, The Relationship of Verbal and Nonverbal Communication, 77-91. The Hague: Mouton.

Firck, R.W. (1985) Communicating emotion: The role of prosodic features. Psychological Bulletin, 97, 412-429.

Frodi, A. (1985) Variations in parental and non parental response to early infant communication. In M. Reiter and T. Filed (eds.), The Psychobiology of Attachment and Separation, 351-367. Orlando, Fla.: Academic Press.

Fromkin, V.A. and Rodman, R. (1974) An Introduction to Language. New York: Holt, Rinehart and Winston.

Furrow, D. (1984) Younf children's use of prosody. Journal of Child Language, 11, 203-213.

Garnica, O. K. (1977) Some prosodic and paralinguistic features of speech to young children. In C.E. Snow and C.A. Ferguson (eds.), Talking to Children: Language input and acquisition, 63-88. Cambridge: Cambridge University Press.

Grice, H.P. (1957) Meaning. Philosophical Review, 68, 377-388

Grieser, D.L. and Kuhl, P.K. (1988) Maternal speech to infants in a tonal language: Support for univeral prosodic features in motherese. Developmetnal Psychology, 24, 14-20.

Habermas, J. (1970) Introductory remarks to a theory of communicative competence. In H.P. Dreitzel (ed.), Recent Sociology, No 2. London: Macmillan.

Halliday, M.A.K. (1973) Explorations in the Functions of Language. London: Edward Arnold.

Halliday, M.A.K. (1975) Learning How to Mean: Explorations in the development of language. London: Edward Arnold.

Halliday, M.A.K. (1978) Meaning and the construction of reality in early childhood. In H.L. Pick and E. Saltzman (eds.). Modes of Perceiving and Processing Information, 67-96. Hillsdale: Erlbaum.

Harding, C.G. and Golinkoff, R. (1979) The origins of intentional vocalisations in prelinguistic infants. Child Development, 50, 33-40.

t' Hart, J., Collier, R. and Cohen, A. (1990) A Perceptual Study of Intonation: An experimental phonetic approach to speech melody. Cambridge: Cambridge University Press.

Haugen, E. (1949) Phoneme or prosodeme. Language, 25, 278-282.

Hess, E. (1970) Ethology and developmental psychology. In P. Mussen, Carmichael's Manual of Child Psychology, 1-38. New York: Wiley.

Hirano, M., Kurita, S. and Nakashima, T. (1981) The structure of the vocal folds. In K. Stevens and M. Hirano (eds.), Vocal Fold Physiology, 33-43. Tokyo: University of Tokyo Press.

Howel, D.C. (1992) Statistical Methods for Psychology. Dusxbury Press.

Hubley, P.A. (1983) The Development of Cooperative Action in Infancy. Unpublished Doctoral Dissertation. The University of Edinburgh.

Hubley, P. and Trevarthen, C. (1979) Sharing a task in infancy. In I. Uzgiris (ed.), Social Interaction during Infancy: New Directions for Child Development, 4, 57-80. San Francisco: Jossey-Bass.

Hutt, S.C., Hutt, C., Lenard, H.G., Bernuth, H.V., and Muntjewerff, W.J. (1968) Auditory responsitivity in the human neonate. Nature, 218, 888-90.

Jakobson, R. and Halle, M. (1957) Phonology in relation to phonetics. In L. Kaiser (ed.), Manual of Phonetics, 215-251. Amsterdam: North Holland Pub. Co.

Jespersen, O. (1922) Language: Its nature, development and origin. London: George Allen and Unwin.

Kaplan, E. and Kaplan, G. (1971) The prelinguistic child. In J. Eliot (ed.), Human Development and Cognitive Processes. New York: Holt, Rinehart and Winston.

Kaye, K. and Fogel, A. (1980) The temporal structure of face-to-face communication between mothers and infants. Developmental Psychology, 16, 454-464.

Keating, P. and Buhr, R. (1978) Fundamental frequency in the speech of infants and children. Journal of the Acoustical Society of America, 63, 567-571.

Keller, H. and Schölmerich, A. (1987) Infant vocalisations and parental reactions during the first four months of life. Developmental Psychology, 23, 62-67.

Keller, E. (1994) Signalyse™: A user's guide.

Kent, R.D. and Miolo, G. (1995) Phonetic abilities in the first year of life. In P. Fletcher and B. MacWhinney (eds.), The Handbook of Child Language, 303-334. Oxford: Blackwell.

Kessen, W., Levine, J. and Wendrich, K. (1979) The imitation of pitch in infants. Infant Behaviour and Development, 2, 93-100.

Klatt, D.H. (1973) Discrimination of fundamental frequency contours in synthetic speech: Implications for models of pitch perception. Journal of the Acoustical Society of America, vol 53, No 1, 8-16.

Kugiumutzakis, G. (1993) Intersubjective vocal imitation in early mother-infant interaction. In J. Nadel and L. Camaioni (eds.), New Perspectives in Early Communicative Development, 23-45. London: Routledge.

Ladd, D. R. and Cutler, A. (1983) Introduction: Models and measurements in the study of prosody. In A. Cutler and D.R. Ladd (eds.), Prosody: Models and measurements, 1-10. Berlin: Springer-Verlag.

Ladd, D.R., Silverman, K.E.A., Tolkmitt, F., Bergmann, G. and K.R. Scherer (1985) Evidence for the independent function of intonation contour type, voice quality, and Fo range in signalling speaker affect. Journal of the Acoustical Society of America, 78 (2), 435-444.

Ladefoged, P. (1962) Elements of Acoustic Phonetics. Chicago: University of Chicago Press.

Laver, J. (1994) Principles of Phonetics. Cambridge: Cambridge University Press.

Lehiste, I. (1970) Suprasegmentals. Cambridge MA: M.I.T. Press.

Lenneberg, E. (1967) Biological Foundations of Language. New York: Wiley Press.

Lewedag, V.L., Oller, K.D. and Lynch, M.P (1994) Infants' vocalisation patterns across home and laboratory environments. First Language, 14, 49-65.

Lewis, M. M. (1936) Infant Speech: A study in the beginnings of language. New York: Harcourt Brace.

Liddell, D and Scott, S. (1882) Dictionary of Ancient Greek Language. Oxford: Oxford Unviversity Press.

Lieberman, P. (1967) Intonation, Perception and Language. Cambridge, Mass: M.I.T. Press.

Lieberman, M. and Sag I. (1974) Prosodic form and discourse function. In Papers from the Tenth Regional Meeting, Chicago Lignuistic Society, 416-426.

Lipourlis, D. (1978) Ancient Greek Prosody (Αρχαια Ελληνικη Μετρικη). Thessaloniki: Observer.

Locke, J.L. (1993) The Child's Path to Spoken Language. Cambridge: Harvard University Press.

Locke, J.L. (1995) Development of the capacity for spoken language. In P. Fletcher and B. MacWhinney (eds.), The Handbook of Child Language, 278-302. Oxford: Blackwell.

Locke, J.L. (1996) Why do infants begin to talk? Language as an unintended consequence. Journal of Child Language, 23, 251-268.

Lorenz, K. (1966) On Aggression. London: Methuen.

Lynch, M.P., Oller, D.K. Steffens, M.L. and Buder, E. (1995) Phrasing in prelinguistic vocalisations. Developmental Psychology, 28, 3-25.

Lynip, A.W. (1951) The use of magnetic devices in the collection and analysis of the preverbal utterances of an infant. Genetic Psychology Monographs, 44, 221-262.

Lyons, J. (1977) Semantics, vol 2. Cambridge: Cambridge University Press.

Malatesta, C.Z. (1981) Infant emotion and the vocal affect lexicon. Motivation and Emotion, 5, 1-23.

Marcos, H. (1987) Communicative functions of pitch range and pitch direction in infants. Journal of Child Language, 14, 255-268.

Marler, P., Evans, C.S. and Hauser, M.D. (1992) Animal signals: Motivational, referential or both?. In H. Papousek, U. Jürgens and M. Papousek (eds.), Nonverbal Vocal Communication: Comparative and developmental approaches, 66-86. Cabridge: Cambridge University Press.

Martinet, A. (1964) Elements of General Linguistics (trans. E. Palmer). London: Faber and Faber.

Marwick, H.M. (1986) The Intonation of Mothers and Children in Early Speech. Unpublished Doctoral Dissertation. The University of Edinburgh.

Masataka, N. (1992) Early ontogeny of vocal behaviour of Japanese infants in response to maternal speech. Child Development, 63, 1177-1185.

Mehler, J., Bertoncini, J., Barriere, M. and Jassik-Gerschenfeld, D. (1978) Infant recognition of mother's voice. Perception, 7, 491-497.

Mehler, J. and Bertoncini (1979) Infant's perception of speech and other acoustic stimuli. In J. Morton and J.C. Marshall (eds.), Psycholinguistics Series-2. London: Elek Science.

Menyuk, P. and Bernholtz, M. (1969) Prosodic features and children's language production. M.I.T. Quarterly Progress Reports, Electronics Research Laboratory, 93, 216-19.

Menyuk, P., Menn, L. and Silber, R. (1979) Early strategies for the perception and production of words and sounds. In P. Fletcher and M. Garman, Language Acquisition, 198-222. Cambridge: Cambridge University Press.

Messer, D. Referential communication: Making sense of the social and physical worlds. In G. Bremner, A. Slater and G. Butterworth, Infant Development: Recent Advances, 291-309. Psychology Press.

Meyer-Eppler, W. (1957) Realisaton of prosodic features in whispered speech. Journal of the Acoustical Society of America, 29, 104-6.

Morse, P.A. (1974) Infant speech perception: A preliminary model and a review of the literature. In R.L. Schiefelbuschand and L. Lloyd (eds.), Language Perspectives: Acquisition, retardation and intervention, 19-53. London: Mcmillan.

Murray, L. (1980) The Sensitivities and Expressive Capacities of Young Infants in Communication with their Mothers. Unpablished Doctoral Dissertation. The University of Edinburgh.

Murray, L. and Trevarthen C. (1985) Emotional regulation of interactions between two-month-olds and their mothers. In T.M. Field and N.A. Fox (eds.), Social Perception in Infants, 177-97. Norwood, NJ: Ablex.

Netsell, R. (1981) The acuqisition of speech motor control: A perspective with direction for research. In R. Stark (ed.), Language Behaviour in Infancy and Childhood, 127-156. Amsterdam: Elsevier-North Holland.

Nettl, B. (1956) Infant musical development and primitive music. Southwestern Journal of Anthropology, 12, 87-91.

Ogden, C.K. and Richards, I.A. (1923) The Meaning of Meaning. London:

Ohala J.J. (1978) The production of tone. In V.A. Fromkin (ed.), Tone: A linguistic survey, 5-39. New York: Academic Press.

Oller, D.K. (1980) The emergence of the sounds of speech in infancy. In G. Yeni-Komshian, J.F. Kavanagh and C.A. Ferguson (eds.), Child Phonology, vol 1, 93-112. London: Academic Press.

Oller, D.K. (1986) Metaphonology and infant vocalisations. In B. Lindblom and R. Zetterstrom (eds.), Precursors of Early Speech, 21-35. New York: Stockton Press.

Ostwald, P.F. (1963) Soundmaking: The acoustic communication of emotion. Springfield, Illinois: Charles C. Thomas.

Papousek, H. (1969) Individual variability in learned responses in human infants. In R.J. Robinson (ed.), Brain and Early Behaviour, 251-266. London: Academic Press.

Papousek, M. (1989) Determinants of responsiveness to infant vocal expression of emotional state. Infant Behaviour and Development, 12, 507-524.

Papousek, M. (1992) Early ontogeny of vocal communication in parent-infant interactions. In H. Papousek, U. Jürgens and M. Papousek (eds.), Nonverbal Vocal Communication: Comparative and developmental approaches, 230-261. Cambridge: Cambridge University Press.

Papousek, H., Jürgens and M. Papousek (1992) Non Verbal Vocal Communication: Comparative and developmental approaches. Cambridge: Cambridge University Press.

Papousek, H. and Papousek, M. (1987) Intuitive parenting: A dilectic counterpart to the infant's integrative competence. In J.D. Osofsky (ed.), Handbook of Infant Development. New York: Wiley.

Papousek, M and Papousek, H. (1981) Musical elements in the infant's vocalisations: Their significance for communication, cognition and creativity. In L.P.

Lipsitt (ed.), Advances in Infancy Research, vol 1, 163-224. Norwood N.J.: Ablex.

Papousek, M., Papousek, H. and Bornstein, M.H. (1985) The naturalistic vocal environment of young infants: On the significance of homogeneity and variability in parental speech. In T. Field and N. Fox (eds.), Social Perception in Infants, 269-297. Norwood N.J.: Ablex.

Papousek, M., Papousek, H. and Harris, B.J. (1984) The emergence of play in parent-infant intercations. In D. Görlitz and J.E. Wohlwill (eds.), Curiosity, Imagination and Play: On the development of spantaneous cognitive and motivational processes, 214-246 Hillsdale: Erlbaum.

Papousek, M., Papousek, H. and Symes, D. (1991) The meaning of melodies in motherese in tone and stress languages. Infant Behaviour and Development, 14, 414-440.

Petrovich-Bartell, N, Cowan, N. and Morse, P.A. (1982) Mother's perceptions of infant distress vocalisations. Journal of Speech and Hearing Research, 25, 371-376.

Piaget, J. (1952) The Origins of Intelligence in Children. New York: International University Press.

Pierhumbert, J. and Hirschberg, J. (1990) The meaning of intonational contours in the interpretation of discourse. In P.R. Cohen, J. Morgan and M.E. Pollack (eds.), Intentions in Communication, 271-311. Cambridge MA: M.I.T. Press.

Pike, K.L. (1945) The intonation of Americal English. In D. Bolinger (ed.), Intonation. Harmondsworth: Penguin.

Plutchik, R. (1981) A general psychoevolutionary theory of emotion

Pollack, I. (1968) Detection of rate of change of auditory frequency. Journal of Experimental Psychology, vol 77, No 4, 535-541.

Reddy, V. (1989) Play with others' expectations: teasing and mucking about in the first year. In A. Whiten (ed.) Natural Theories of Mind, 143-158. Oxford: Balckwell.

Reinholt Petersen, N. (1986) Perceptual compensation for segmentally conditioned fundamental frequency perturbation, Phonetica, 43, 31-42.

Robb, M.P and Saxman, J.H. Developmental trends in vocal fundamental frequency of young children. Journal of Speech and Hearing Research, 28, 421-427.

Roederer, J.G. (1995) The Physics and Psychophysics of Music: An introduction. Springer-Verlag: New York.

Rosenblith, W.A. and Stevens K.N. (1953) On the DL for frequency. Journal of the Acoustical Society of America, 25, 980-985.

Rossing, A. (1989) The Science of Sound Processing. Cambridge: Cambridge University Press.

Ryan, J. (1974) Early language development. In M.P.M Richards (ed.), The Integration of the Child into a Social World, 185-213. Cambridge: Cambridge University Press.

Sachs, J., Brown Rt. and Salerno, R.A. (1976) Adults' speech to children. In W. von Raffler-Engel and Y. Lebrun (eds.), Baby Talk and Infant Speech. Lisse, Netherlands: Swets and Zeitlinger.

Scherer, K.R. (1982) Methods of research on vocal communication: Paradigms and parameters. In K.R. Scherer and P. Ekman (eds.), Handbook of Methods in Nonverbal Behaviour Research. Cambridge: C.U.P.

Scherer, K.R. (1984) The assessment of vocal expression in infants and children. In C.E. Izard (ed.), Measuring Emotions in Infants and Children, 127-163, Cambridge MA: Cambeidge University Press.

Searle, J. (1969) Speech Acts. Cambridge: Cambridge University Press.

Seashore, C.E. (1967) Psychology of Music. New York: Dover.

Sheppard, W.C. and Lane, H.L. (1968) Development of the prosodic features of infant vocalisation. Journal of Speech and Hearing Research, 11, 94-108.

Shimura, Y., Imaizumi, S. and Shutou, T. (1992) Perceptual evaluation of emotional aspects of infants' vocalisations. University of Tokyo Annual Bulletin, 3, 87-96.

Shimura, Y. and Imaizumi, S. (1995) Emotional information in young infants' vocalisations. Proceedings ICPhS, vol 3, 412-415.

Siegel, G.M., Cooper, M., Morgan, J.L. and Brenneise-Sarshad, R. (1990) Imitation of intonation by infants. Journal of Speech and Hearing Research, 33, 9-15.

Small, A.M. (1973b) Psychoacoustics. In F.D. Minifie, T.J. Hixon and F. Williams (eds.), Normal Aspects of Speech, Hearing and Language. Englewood Cliffs, N.J.: Prentice-Hall.

Snow, C.E. (1977) The development of conversation between mothers and babies. Journal of Child Language, 4, 1-22.

Spencer, H. (1858) Essays Scientific, Political and Speculative. London: Longman.

Stark, R.E. (1980) Stages of speech development in the first year of life. In G. Yeni-Komshian, J.F. Kavanagh and C.A. Ferguson (eds.), Child Phonology, vol 1, 73-92. London: Academic Press.

Stark, R.E., Rose, S.N. and McLagen, M. (1975) Features of infant sounds: The first eight weeks of life. Journal of Child Language, 2, 205-222.

Steele, J. (1779) Prosodia Rationalis. London: Nichols

Stern, D.N. (1985) The Interpersonal World of the Infant: A view form psychoanalysis and development psychology. New York: Basic Books.

Stern, D.N., Hoffer, L., Haft W., and Dore J. (1985) Affect attunement: The sharing of feeling states between mother and infant by means of inter-modal fluency. In T. Field and N. Fox, Social Perception in Infants, 249-268.

Stern, Koslowski (1979) The origins of reciprocity: The early mother-infant interaction. In M. Lewis and L.A. Rosenblum (eds.), The Effect of the Infant on its Caregiver, 49-76. New York: WIley.

Stern, D.N., Spieker, S. Barnett, R.K. and Mckain, K. (1983) The prosody of maternal speech: Infant age and context related changes. Journal of Child Language, 10, 1-15.

Studderd-Kennedy, M. (1986) Sources of variability in early speech development. In J.S. Perkell and D.H. Klatt, Invariance and Variability in Speech Processes, 58-76. Hillsdale, N.J. Erlbaum.

Taine, H. (1877) Acquisition of language by children. Mind, 2, 252-259.

Tomasello, M. (1993) On the interpersonal origins of self concept. In U. Neisser (ed.), Ecological and Interpersonal Knowledge of the Self, 174-184. New York: Cambridge University Press.

Tomasello, M. and Farrar, J. (1986b) Joint attention and early language. Child Development, 57, 1454-1463.

Tonkova-Yampols'kaya, R.V. (1969) Development of speech intonation in infants during the first two years of life. Soviet Psychology, 7, 48-54.

Trehub, S.E., Trainor, L.J. and Unyk, A.M. (1993) Music and speech processing in the first year of life. Advances in Child Development and Behaviour, vol 24, 2-34.

Trevarthen, C. (1982) The primary motives for cooperative understanding. In G. Butterworth and P. Light (eds.), Social Cognition: Studies of the development of understanding, 77-109. Brighton: Harvester Press.

Trevarthen, C. (1986a) Development of intersubjective motor control in infants. In M.G. Wade and H.T.A. Whiting (eds.), Motor Development in Children: Aspects of Coordination and Control, 209-61. Dordrecht: Martinus Nijhof.

Trevarthen, C. (1987) Sharing makes sense: Intersubjectivity and the making of an infant's meaning. In R. Steele and T. Treadgold, Language Topics, 177-208. Amsterdam: J. Benjamins.

Trevarthen, C. (1990) Signs before speech. In T.A. Sebeok and J. Umiker-Sebeok (eds.), The Semiotic Web, 689-755. Berlin: Mouton de Gruyter.

Trevarthen, C. (1993) Infant semiosis. In W. Nöth (ed.) Origins of Semiosis, 1-12. Berlin: Mouton de Gruyter.

Trevarthen, C. and Hubley, P. (1978) Secondary intersubjectivity: Confidence, confiding and acts of meaning in the first year. In A. Lock (ed.), Action: Gesture and symbol: The emergence of language, London: Academic Press.

Trevarthen, C. and Malloch, S. (1997) Analysis of Maternal Speech to Infants. Research Project.

Trevarthen, C. and Marwick, H.M. (1986) Signs of motivation for speech in infants and the nature of a mother's support for development of language. In B. Lindblom and R. Zetterstrom (eds.), Precursors of Early Speech, 279-308. New York: Stockton Press.

Tsumura, T. (1976) Auditory thresholds of frequency change in brief tones. Journal of the Acoustical Society of Japan, 32, 377-386.

Tsumura, T., Sone, T. and Nimura, D. (1973) Auditory detection of frequency transition. Journal of the Acoustical Society of America, vol 53, No 1, 17-25.

Tsumura, T., Nakajima, Y. and Taranishi, R. (1990) Discrimination of rate of frequency change and dynamic pitch perception. Journal of the Acoustic Society of Japan, vol 46, No 5, 403-412.

Vihman, M.M., Mcken, M.A., Miller, R., Simmons, H. and Miller, J. (1985) From babbling to speech: A reassessement of the continuity issue. Language, 61, 395-443.

Vihman, M.M., Ferguson, C.A. and Elbert, M. (1986) Phonological development from babbling to speech: Common tendencies and individual differences. Applied Psycholinguistics, 7, 3-40.

Weeks, T.E. (1982) Intonation as an early marker of meaning. In M.R. Key (ed.), Nonverbal Communication Today, 157-168. Berlin: Mouton.

Weinbergh, M.K. and Tronick, E.Z. (1994) Beyond the face: An empirical study of infant affect configurations of facial, vocal, gestural and regulatory behaviours. Child Development, 65, 1503-1515.

Weir, R. (1962) Language in the Crib. The Hague: Mouton.

Welmers, W. E. (1959) Tonemics, morphotonemics, and tonal morphemes. Gen. Linguistics, 4, 1-9.

Werner, H. and Kaplan, B. (1963) Symbol Formation: An organismic-developmental approach. New York: Wiley.

Whalen, D.H., Levitt, A.G., Hsiao, P.L. and SMorodinsky, I. (1995) Intrinsic Fo of vowels in the babbling of 6-,9-, and 12-month-old French- and English-learning infants. Journal of the Acoustical Society of America, 97 (4), 2533-2539.

White, S.J. (1982) Nonverbal antecendents to language functioning: A model and its relevance for the deaf. In M.R. Key (ed.), Nonverbal Communication Today, 234-243. Berlin: Mouton.

Wiener, M., Devoe, S., Rubinow, S. and Geller, J. (1972) Non verbal behaviour nad non verbal communication. Psychological Review, vol 79, No 3, 185-214.

Wolff, P.H. (1969) The natural history of crying and other vocalisations in infancy. In B.M. Foss (ed.), Determinants of Infant Behaviour IV, 81-109. London: Methuen.

Wong, J. and Miller, C. (1984) Parental Oerceotion of Infant Vocal Development. Poster presented in the International Conference on Infant Studies. New York: NY.

Young, G. and Décarie, T.G. (1977) An ethology based catalogue of facial/vocal behaviour in infancy. Animal Behaviour, 25, 95-107.

Zelazo, P.R., Kagan, J. and Hartmann, R. (1975) Excitement and boredom as determiantns of vocalisation in infants. Journal of Genetic Psychology, 126, 107-117

Zinober, B. and Martlew, M. (1985) The development of communicative gestures. In M. Barrett, Children's Single-Word Speech, 183-210. New York: Wiley.

Zlatin-Leufer, M.Z. and Horii, Y. (1977) Fundamental frequency characteristics of infant non-distress vocalisation during the first twenty four weeks. Journal of Child Language, 4, 171-184.

APPENDIX I

CODING SYSTEM FOR THE VIDEO RECORDS

I. Intersubjective Functions

INTERPERSONAL

a) Response to Invitation to Play

1. *Looks at Mother's Face:* The infant clearly looks at mother's face or eyes when she is acting playfully to him or her.

2. *Looks at Mother's Body:* The infant looks at mother's body when she is acting playfully to him or her.

3. *Touches Head:* The infant playfully touches the mother's face or head in response to her playful actions.

4. *Touches Body:* The infant playfully touches the mother's body in response to her playful actions.

5. *Dances:* The infant bounces or claps hands in response to the mother's singing, imitating the rhythm of the song

6. *Imitates Facial Expression:* Immediately following the mother, the infant matches her face expression with a similar action with own face.

7. *Imitates Body Movement:* Immediately following the mother, the infant matches a body or hand movement she has made with a similar action.

8. *Complies in Interpersonal Play:* The infant acts according to the mother's directives in a playful activity immediately after the directive is given e.g. claps hands when asked to do so, following the ritual of a nursery action rhyme.

9. *Accepts Assistance:* The infant spontaneously adopts the appropriate posture to allow the mother to perform a playful action, or performs an appropriate action to complete the mother's activity.

10. *Shows Affection*: The infant hugs mother or nuzzles into her as a response to her similar expressions.

b) Invitation to Mother to Play

11. *Touches Head*: The infant acts playfully on the mother's face and head e.g. grabs mother's face, pulls her hair.

12. *Touches Body*: The infant acts playfully on the mother's body e.g. bites her finger.

13. *Imposes Playful Action*: The infant moves mother's hand to perform a playful action e.g. claps mother's hand together.

14. *Teases by Disobeying*: The infant pretends to disobey a prohibition, or to fulfil the other's expectation and then calls off with amusement.

15. *Teases by Resisting*: Resisting mother's actions or intentions with amusement.

c) Attention Seeking

16. *Looks at Unresponsive Mother's Face:* The infant looks at the mother's face or eyes, while the mother is unresponsive.

17. *Looks at Unresponsive Mother's Body:* The infant looks at the mother's body, while the mother is unresponsive.

18. *Seeks for Mother:* The infant looks where the mother left in the condition 'Infant Alone'.

19. *Acts on Unresponsive Mother's Face:* The infant acts on mother's face, while she is unresponsive.

20. *Acts on Unresponsive Mother's Face:* The infant acts on mother's body, while she is unresponsive.

21. *Reaches at Unresponsive Mother:* The infant reaches towards the mother, while she is unresponsive.

22. *Hugs Unresponsive Mother:* The infant hugs mother or nuzzles into her, while she is unresponsive.

23. *Searches for Mother:* The infant moves towards the place the mother left in the condition 'Infant Alone'.

d) Invitation to Stranger

24. *Looks at Researcher's Face:* The infant looks at the researcher's face or eyes.

25. *Looks at Researcher's Body:* The infant looks at the researcher's body.

26. *Reaches at Researcher:* The infant reaches towards the researcher.

e) Request for Action on Self

27. The infant extends its hand upwards towards the mother.

f) Conflict

28. *Withdraws*: The infant pulls own hand or face away from touching or imposing by the mother.

29. *Resists*: The infant pushes away the mother's hand or makes other attempts against imposition or touching.

g) Showing Off

30. The infant performs exaggerated expressive actions for the entertainment of the mother, with or without an object. These performances call for no response from the partner other than an expression of appreciation by watching, smiling etc.

REFERENTIAL

h) Converging Interest

31. *Shows Interest:* Looks at The infant looks at the object the mother is handling or has been handling immediately before.

32. *Shows Interest in Action:* The infant looks at the mother when performing an action with an object.

33. *Takes*: After the mother has offered an object, the infant takes it.

34. *Gives*: The infant hands an object to the mother after she has asked for it.

35. *Imitates Praxic Action*: Immediately following an action the mother has performed with an object(s), the infant performs a similar action.

36. *Imitates Imaginative Action*: Immediately following the mother's imaginative action with an object(s) (or an imagined object(s), the infant performs a similar action.

37. *Follows in Manipulation*: The infant grasps, touches or manipulates an object which the mother is manipulating or has been manipulating immediately before.

38. *Shows Interest in Object*: The infant reaches to an object the mother is handling or has been handling immediately before.

39. *Shares Interest in Object*: The infant touches mother's hand when she is touching an object.

40. *Reaches at Object:* The infant's palm(s) is(are) open in a grasp position in front of the object the mother is manipulating or has been manipulating immediately before.

41. *Takes Spontaneously*: The infant removes, or tries to remove, object from mother's hand without it being offered.

i) Joint Interest

42. *Follows Gaze*: The infant looks in the same direction as the mother.

43. *Follows Point*: The infant looks in the direction indicated by the mother's pointing or other directive action (giving, offering, showing).

44. *Complies*: The infant acts according to the mother's directive immediately following the directive.

45. *Acquiesces*: The infant contributes to an action the mother is trying to impose or perform.

46. *Completes Action*: The infant completes an action in play with an object which the mother is assisting. The infant also interrupts his/her activity while the mother helps.

47. *Accepts Assistance*: The infant accepts without protest the mother's assistance in a situation where the infant is having difficulties to complete an action on an object. The infant interrupts his/her activity while the mother helps.

j) Directives

48. *Alternates Gaze*: The infant alternates gaze from an object that he/she is observing with the to the mother and back at least once.

49. *Offers*: The infant holds out an object to the mother and permits her to take it.

50. *Shows*: The infant holds out an object to the mother but does not permit her to take it

51. *Gives Spontaneously*: The infant puts an object into the mother's hand or mouth without request.

52. *Points*: With palm or index finger the infant points object.

53. *Touches with Object*: The infant touches part of the mother's body with an object.

k) Request for Object

54. The infant extends his or her hands towards an object the mother is now holding, which she has taken immediately before from the baby without this being offered.

l) Conflict on Object

55. The infant violently pulls an object the mother is holding away from her, while the mother is not letting him or her having it.

m) Split Attention

56. A situation is codified as such when a vocalisation occurs while the infant is shifting his or her attention from an activity on an object with the mother to another object.

II. Solitary Functions

INFANT - OBJECT

57. *Looks at Object:* The infant concentrates his/her gaze on a target (a manipulable object or an event in the environment).

58. *Looks at Moving Object:* The infant tracks a moving object.

59. *Scans:* The infant glances around the room without focusing on any particular object.

60. *Spots Object:* As the infant scans around the room, he/she notices a target and fixes gazes on it.

61. *Listens:* The infant attends to a sound in the environment (cocking head).

62. *Holds Object:* The infant holds a toy(s) in hand(s).

63. *Inspects Object:* The infant attempts to modify an object's orientation or configuration by displacing it with the hand while looking at it.

64. *Manipulates Object:* The infant handles a toy in way so as to make it operate.

65. *Explores Object:* The infant bangs, bounces, swings, rolls, shakes, pushes or pulls.

66. *Reaches Object*: The infant leans or moves towards an object.

67. *Moves Rhythmically:* The infant moves his/her arms or legs, or bounces in the rhythm of the tune played by the music box.

III. Emotions

68. *Shared Positive Emotion*: The infant smiles in a Mother-Infant or Mother-Infant-Object situation. The smile is recognised intuitively. No attempt is made to define the facial movements.

69. *Shared Negative Emotion*: The infant makes distressed face in a Mother-Infant or Mother-Infant- Object situation. The smile is recognised intuitively. No attempt is made to define the facial movements.

70. *Non-Shared Positive Emotion*: The infant smiles in an Infant-Object situation. The smile is recognised intuitively. No attempt is made to define the facial movements.

71. *Non-Shared Negative Emotion*: The infant makes distressed face in a Infant-Object situation. The distressed face is recognised intuitively. No attempt is made to define the facial movements.

APPENDIX II

TABLES

Table 5.1: Robin, 7 months. Distribution of Messages by Contour Shape. Percentages of contour shapes are shown in parentheses.

MESSAGE	CONTOUR SHAPE						TOTALS
	Level	Fall	Undulating Fall	Bell	Rise	Undulating Rise	
Invitation to Stranger	1 (12.5)				7 (87.5)		8
Non Shared Interest	1 (16.7)				5 (83.3)		6
Shared Interest				2 (13.4)	13 (86.6)		15
Attention Seeking		1 (10)			9 (90)		10
Request		1 (25)		1 (25)	2 (50)		4
Shared Negative Emotion		2 (7)	7 (24.1)	6 (20.6)	4 (13.8)	10 (34.5)	29

Table 5.2: Robin, 7 months. Prosodic patterns of different messages and their characteristics.

PATTERN	MESSAGE	PROSODIC FEATURES and STATISTICS									
		Peak Pitch (semitones)		Mean Pitch (semitones)		Duration (milliseconds)		Pitch Range (semitones)		Rate of Pitch Change (semitones/sec)	
A	Non-Shared Interest	102.89 (100.45-105.38)	F(5,67)=5.6 p<.001	102.04 (100.05-104.92)	F(5,67)=3.53 p<.01	250 (138-575)	F(5,67)=22.14 p<.001	1.72 (0.7-6.34)	F(5,67)=4.78 p<.001	11.85 (10.42-14.98)	F(5,58)=4.41 p<.01
	Shared Interest	104.86 (100.18-111.03)		104.06 (99.429-110.55)		290 (225-530)		1.8 (1-5.11)		13.1 (7.54-27.18)	
	Invitations to Stranger	103.01 (99.809-106.52)		101.63 (99.79-105.47)		300 (155-479)		1.95 (0.05-4.23)		13.1 (9.88-19.12)	
B	Attention Seeking	105.69 (104.2-111.86)		104.31 (103.04-109.97)		290 (80-870)	ibd	4.24 (0.34-6)		32.88 (13.78-45)	
C	Shared Negative Emotion	106.81 (103.73-119.38)		105.23 (101.55-116.18)		1115 (345-4480)		4.46 (1.27-10.38)		13.97 (7.16-52.69)	
	Requests	106.23 (103.73-107.17)		105.4 (102.99-105.59)		820 (575-2274)		4.22 (3.13-5.58)		20.95 (16.92-24)	ibd

Table 5.3: Robin, 7 months. Correspondence between Messages attributed by the Mother and Video Analysis.

VIDEO ANALYSIS		MESSAGE						TOTALS
		Non-Shared Interest	Shared Interest	Invitation to Stranger	Attention Seeking	Request	Shared Negative Emotion	
Person-Person	Attention Seeking				9 (90)			9
	Response to Invitation to Play					1 (25)	2 (6.9)	3
	Conflict						5 (17.2)	5
	Invitation to Stranger			8 (100)				8
Mother-Infant-Object	Converging Interest		13 (86.7)			3 (75)	22 (75.9)	38
	Joint Interest		2 (13.3)					2
	Split Attention	1 (16.7)						1
Infant-Object		5 (83.3)			1 (10)			6

Table 5.4: Robin, 8 months. Distribution of Messages by Contour Shape. Percentages of contour shapes are shown in parentheses.

MESSAGE	CONTOUR SHAPE						TOTALS
	Level	Fall	Undulating Fall	Bell	Rise	Undulating Rise	
Invitation to Stranger	1 (10)				8 (80)	1 (10)	10
Shared Interest				1 (14.3)	6 (85.7)		7
Request		1 (8.3)			10 (83.4)	1 (8.3)	12
Shared Negative Emotion		2 (12.5)		2 (12.5)		12 (75)	16
Shared Positive Emotion	1 (12.5)			3 (37.5)	4 (50)		8

Table 5.5: Robin, 8 months. Prosodic patterns of different messages and their characteristics.

PATTERN	MESSAGE	PROSODIC FEATURES and STATISTICS										
		Peak Pitch (semitones)			Mean Pitch (semitones)		Beginning Pitch (semitones)		Final Pitch (semitones)			Duration (milliseconds)	
*	Shared Interest	106.66 (104.98-110.22)	F(4,49)=6.98 p<.001		104.23 (103.66-109.5)		102.28 (101.71-108.79)		106.66 (104.11-110.1)	F(4,49)=3.75 p<.01		330 (252-522)	F(4.49)=14.08 p<.001
A	Invitations to Stranger	104.96 (101.51-107.93)		F(3,41)=3.4, p<.05 F(2,30)=3.77, p<.05	103.64 (101.58-107.23)	F(4,49)=5.2 p<.001	100.98 (98.86-106.66)	F(4.49)=6.66 p<.001	104.78 (101.32-107.93		F(3,41)=3.87 p<.05	310 (167-635)	
B	Requests	107.78 (100.76-113)			105.15 (100.84-105.93)	ibd	101.33 (99.14-105.81)	ibd	108.17 (100.76-112.86)			360 (146-610)	ibd
C	Shared Negative Emotion	106.48 (104.9-110.86)	ibd		104.63 (102.24-108)	ibd	102.22 (99.42-106.52)	ibd	105.33 (102.09-110.4)	ibd	ibd	1070 (307-2744)	
D	Shared Positive Emotion	119.83 (105.23-114.72)			107.69 (103.37-113.13)		106.9 (102-110.72)		108.33 (102.66-113.22)			395 (321-1673)	ibd

Table 5.6: Robin, 8 months. Correspondence between Messages attributed by the Mother and Video Analysis.

VIDEO ANALYSIS		MESSAGE					TOTALS
		Ivnitation to Stranger	Shared Interest	Request	Shared Positive Emotion	Shared Negative Emotion	
Person-Person	Attention Seeking					3 (18.75)	3
	Conflict		2 (28.6)			10 (62.5)	12
	Invitation to Stranger	10 (100)			1 (12.5)		11
Mother-Infant-Object	Converging Interest		5 (71.4)	12 (100)	6 (75)	3 (18.75)	26
	Directive				1 (12.5)		1

Table 5.7: Robin, 9 months. Distribution of Messages by Contour Shape. Percentages of contour shapes are shown in parentheses.

MESSAGE	CONTOUR SHAPE						TOTALS
	Level	Fall	Undulating Fall	Bell	Rise	Undulating Rise	
Invitation to Stranger	1 (6.25)	1 (6.25)	1 (6.25)	1 (6.25)	10 (62.5)	2 (12.5)	16
Non-Shared Interest				1 (16.6)	4 (66.8)	1 (16.6)	6
Shared Interest	1 (8.3)		1 (8.3)	1 (8.3)	9 (75.1)		12
Attention Seeking		1 (16.6)		1 (16.6)	3 (50)	1 (16.6)	6
Shared Negative Emotion				1 (20)		4 (80)	5
Request				5 (62.5)	1 (12.5)	2 (25)	8
Shared Comment		3 (60)	1 (20)		1 (20)		5
Shared Narrative		1 (25)	2 (50)	1 (25)			4
Shared Positive Emotion		3 (25)		4 (33.3)		5 (41.7)	12

Table 5.8: Robin, 9 months. Prosodic patterns of different messages and their characteristics.

PATTERN	MESSAGE	PROSODIC FEATURES and STATISTICS									
		Mean Pitch (semitones)			Final Pitch (semitones)			Duration (milliseconds)			Contour Tail
A	Shared Comments	99.93 (97.93-101.19)	F(8,66)=3.41 p<.01		99.14 (96.73-101.51)	F(8,66)=4.67 p<.001	F(7,60)=3.08 p<.01 F(5,41)=2.8 p<.05 F(4,36)=2.97 p<.05	200 (174-412)	F(8,66)=7.2 p<.001	F(7,60)=4.37 p<.001	fall
B	Shared Narratives	102.03 (101.76-102.19)			101.81 (99.86-102.66)	ibd		2000 (696-10055)			fall
	Requests	102.99 (97.42-106.57)			102.09 (97.36-106.88)	ibd		750 (358-1130)			fall
C	Shared Negative Emotion	107.14 (100.09-107.82)			109.05 (102.38-109.77)			600 (345-1723)			rise
*	Invitations to Stranger	103.22 (99.33-107.19)			103.74 (98.57-113.95)			320 (246-1093)	ibd	ibd	rise
	Non-Shared Interest	103.67 (99.75-109.58)			103.29 (98.57-109.61)			270 (237-866)			rise
	Shared Interest	101.79 (99.67-105.65)			101.61 (98.97-105.57)	ibd		270 (158-777)		ibd	rise
	Attention Seeking	103.13 (101.41-106.82)			103.56 (102-107.03)			320 (273-598)		ibd	rise
*	Shared Positive Emotion	100.35 (98.33-104.07)	ibd	F(6,55)=3.27 p<.01	100.84 (97.91-102.97)	ibd	ibd	630 (244-927)			-------

Table 5.9: Robin, 9 months. Correspondence between Messages attributed by the Mother and Video Analysis.

VIDEO ANALYSIS		MESSAGE									TOTALS
		Non-Shared Interest	Invitation to Stranger	Shared Interest	Attention Seeking	Requests	Shared Positive Emotion	Shared Narrative	Shared Comment	Shared Negative Emotion	
Person-Person	Attention Seeking				6 (100)					1 (20)	7
	Response to Invitation to Play					5 (62.5)	4 (33.3)	2 (50)	1 (20)		12
	Conflict									1 (20)	1
	Request for Action on Self					1 (12.5)					1
	Invitation to Stranger		16 (100)				3 (25)				19
Mother-Infant-Object	Converging Interest			12 (100)		2 (25)	3 (25)	2 (50)	4 (80)	3 (60)	26
	Joint Interest						2 (16.7)				2
Infant-Object		6 (100)									6

Table 5.10: Robin, 10 months. Distribution of Messages by Contour Shape. Percentages of contour shapes are shown in parentheses.

MESSAGE	CONTOUR SHAPE						TOTALS
	Level	Fall	Undulating Fall	Bell	Rise	Undulating Rise	
Shared Recognition of Object		1 (10)		3 (80)		1 (10)	5
Request				5 (100)			5
Attention Seeking				3 (75)		1 (25)	4
Shared Negative Emotion		9 (60)	1 (6.7)	1 (6.7)	2 (13.3)	2 (13.3)	15
Enjoyment	1 (7.7)	1 (7.6)	5 (38.5)	5 (38.5)		1 (7.7)	13
Excitement	1 (7.1)	2 (14.3)	6 (42.9)	4 (28.6)		1 (7.1)	14
Shared Interest		1 (25)		2 (50)	1 (25)		4
Shared Comment	2 (25)	1 (12.5)		3 (37.5)	1 (12.5)	1 (12.5)	8

Table 5.11: Robin, 10 months. Prosodic patterns of different messages and their characteristics.

MESSAGE	PROSODIC FEATURES and STATISTICS															
	Peak Pitch (semitones)		Mean Pitch (semitones)		Beginning Pitch (semitones)		Final Pitch (semitones)		Duration (milliseconds)		Pitch Range (semitones)		S.D. of Pitch (semitones)		Rate of Pitch Change (semitones/sec)	
Shared Comments	103.6 (100.71-110.22)	F(7,61)=7.44 p<.001	102.22 (100.24-103.73)	F(7,61)=6.2 p<.001	100.02 (99.75-104.45)	F(7,61)=14.24 p<.001	102.79 (97.55-104.32)		263 (116-599)	F(7,61)=2.47 p<.05	2.27 (0.95-8.41)	F(7,61)=3.94 p<.001	0.68 (0.3-1.09)	F(7,61)=2.44 p<.05	9.71 (5.48-32)	
Shared Interest	103.32 (99.7-112.5)		102.54 (98.45-111.7)		101.46 (96.92-110.69)	ibd	103.03 (98.97-110.69)		339 (253-607)		2.23 (1.81-2.91)	ibd	0.74 (0.5-1.03)		11.77 (6.76-25.26)	
Attention Seeking	102.98 (101.27-104.98)	ibd	101.24 (99.76-103.41)		98.63 (97.113-98.97)	ibd	99.91 (99.42-101.36)		501 (355-751)		4.98 (2.86-6)		1.4 (0.75-1.52)		42.31 (33.49-46.19)	
Requests	103.6 (99.64-104.57)	ibd	102.02 (98.93-104.02)		99.59 (98.33-102.09)	ibd	99.2 (98.51-103.6)	F(7,61)=3.95 p<.001	650 (390-775)		3.1 (1.31-5.11)		1 (0.46-1.43)		26.1 (11.75-43.6)	
Shared Recognition of Object	101.31 (99.75-102.75)	ibd	100.34 (99.09-101.37)		99.92 (98.74-101.51)	ibd	100.45 (98.27-100.96)	ibd	581 (476-764)		2.17 (1.92-3.55)		0.55 (0.5-1.09)		6.35 (3.34-23.41)	F(7,61)=3.4 p<.01
Enjoyment	102.09 (98.8-106.66)	ibd	101.29 (98.22-105.79)		99.97 (96.72-106.52)	ibd	100.02 (97.67-105.57)	ibd	738 (167-1971)		2.24 (1.04-4.98)	ibd	0.61 (0.29-1.82)		16.4 (3.11-62.33)	
Excitement	104.28 (100.24-106.66)	ibd	102.41 (99.82-104.91)		100.12 (98.33-103.06)	ibd	101.04 (97.79-105.92)	ibd	863 (216-1366)		4.97 (1.44-8.37)		1.08 (0.13-1.86)		32.73 (12.99-51.73)	
Shared Negative Emotion	107.8 (103.46-112)		106.35 (100.97-108.33)		107.55 (101.27-111.84)		104.57 (98.45-111.8)	ibd	428 (138-2270)		5.05 (0.93-7.94)		1.42 (0.4-3.35)		22.57 (7.36-55)	

Table 5.12: Robin, 10 months. Correspondence between Messages attributed by the Mother and Video Analysis.

VIDEO ANALYSIS		MESSAGE								TOTALS
		Shared Recognition of Object	Shared Comment	Request	Shared Interest	Attention Seeking	Enjoyment	Excitement	Shared Negative Emotion	
Person-Person	Attention Seeking					4 (100)			3 (20)	7
	Response to Invitation to Play						6 (46.1)	7 (50)		13
	Conflict								9 (60)	9
	Invitation to Stranger		5 (62.5)							5
Mother-Infant-Object	Converging Interest	2 (40)	3 (37.5)	2 (40)	3 (75)		2 (15.4)	5 (35.7)	3 (20)	20
	Joint Interest	2 (40)		3 (60)	1 (25)		4 (30.8)	2 (14.3)		12
	Directive	1 (40)					1 (7.7)			2

Table 5.13: Robin, 11 months. Distribution of Messages by Contour Shape. Percentages of contour shapes are shown in parentheses.

MESSAGE	CONTOUR SHAPE						TOTALS
	Level	Fall	Undulating fall	Bell	Rise	Undulating rise	
Attention Seeking	1 (14.3)	1 (14.3)			5 (71.4)		7
Invitation to Stranger	1 (16.6)				4 (66.8)	1 (16.6)	6
Shared Interest	1 (16.6)				4 (66.8)	1 (16.6)	6
Practising	1 (10)	2 (20)	7 (70)				10
Shared Positive Emotion	1 (6.25)	1 (6.25)	2 (12.5)	10 (62.5)	1 (6.25)	1 (6.25)	16
Shared Negative Emotion	1 (10)	4 (40)	1 (10)		2 (20)	2 (20)	10

Table 5.14: Robin, 11 months. Prosodic patterns of different messages and their characteristics.

PATTERNS	MESSAGE	PROSODIC FEATURES and STATISTICS																	
		Peak Pitch (semitones)			Mean Pitch (semitones)			Final Pitch (semitones)			Duration (milliseconds)			Pitch Range (semitones)			S.D. of Pitch (semitones)		Contour Tail
A	Attention Seeking	101 (99.48-102.38)	F(6,53)=6.5 p<.001	F(4,27)=5.14 p<.01	100.35 (99.02-101.85)	F(6,53)=5.36 p<.001	F(4,27)=3.57 p<.05	100.45 (99.48-102.38)	F(6,53)=6.17 p<.001		240 (174-297)	F(6,53)=7.88 p<.001	F(4,33)=8.27 p<.001	0.79 (0.39-1.84)	F(6,53)=6.19 p<.001	F(6,53)=6.19 p<.001	0.22 (0.08-0.68)	F(6,53)=3.91 P<.01	rise
	Shared Interest	102.13 (100.86-105.34)	ibd		101.63 (100.41-107.29)			102.04 (100.18-108.47)			260 (188-359)	ibd	ibd	1.61 (0.45-3.25)	ibd		0.55 (0.14-1.04)		rise
B	Shared Negative Emotion	108.65 (105.92-111.84)			106.28 (104.32-108.76)			106.29 (101.27-108.63)			760 (342-5058)			6.51 (3.02-8.63)			1.41 (0.75-2.31)		no pattern
C	Practising	104.9 (100.5-106.66)			103.05 (99.98-103.42)			99.83 (97.55-102.38)	ibd	F(6,53)=6.17 p<.001	2300 (149-3853)			6.12 (1.07-8.5)			1.3 (0.38-2.45)		fall
*	Invitations to Stranger	106 (101.06-110.4)			104 (100.97-109.1)			104.3 (100.45-109.61)			670 (237-1767)			3.4 (0.2-6.77)			1.34 (0.08-1.63)		rise
*	Shared Positive Emotion	106.1 (102.28-121.42)			103.97 (102.17-117.1)			103.13 (99.7-112.34)			590 (171-1084)	ibd		4.65 (0.28-14.53)			4.18 (0.12-4.72)		fall

Table 5.15: Robin, 11 months. Correspondence between Messages attributed by the Mother and Video Analysis.

VIDEO ANALYSIS		MESSAGE						TOTALS
		Attention Seeking	Invitation to Stranger	Shared Interest	Shared Positive Emotion	Shared Negative Emotion	Practising	
Person-Person	Attention Seeking	7 (100)				2 (20)		9
	Response to Invitation to Play				5 (31.25)			5
	Conflict					6 (60)		6
	Invites Play				1 (6.25)		7 (70)	8
	Invitation to Stranger		6 (100)					6
Mother-Infant-Object	Converging Interest			2 (33.3)		1 (10)	1 (10)	4
	Joint Interest			4 (66.7)	10 (62.5)			14
	Directive					1 (10)	2 (20)	7

Table 5.17: Julie, 7 months. Distribution of Messages by Contour Shape. Percentages of Contour Shapes are shown in parentheses.

MESSAGE	CONTOUR SHAPE						TOTALS
	Level	Fall	Undulating Fall	Bell	Rise	Undulating Rise	
Non-Shared Narrative		2 (28.6)	1 (14.3)	4 (57.1)			7
Request	2 (50)			2 (50)			4
Shared Negative Emotion		2 (16.7)	2 (16.7)	7 (58.3)	1 (8.3)		12
Shared Narrative		3 (10.3)	5 (17.3)	17 (58.7)		4 (13.7)	29
Shared Positive Emotion		1 (8.3)	1 (8.3)	8 (66.8)	1 (8.3)	1 (8.3)	12
Squeal		2 (15.4)		11 (84.6)			13

Table 5.18: Julie, 7 months. Prosodic patterns of different messages and their characteristics.

PATTERN	MESSAGE	PROSODIC FEATURES and STATISTICS																			
		Peak Pitch (semitones)			Mean Pitch (semitones)			Beginning Pitch (semitones)			Duration (milliseconds)		Pitch Range (semitones)			S.D. of Pitch (semitones)			Rate of Pitch Change (semitones/sec)	
A	Requests	101.81 (101.51-105.22)	F(5,72)=27.92 p<.001	F(5,61)=10.71, p<.001 F(2.37)=7.98, p<.001	101.38 (100.74-104.51)	F(5,72)=27.55 p<.001	F(5,72)=27.55, p<.001 F(2,37)=3.97, p<.05	101.02 (98.51-105.22)	F(5,72)=1.5 p<		F(5,61)=3.63 p<.01	381 (198-570)	F(5,72)=5.34 p<.001	1.47 (0.49-2.9)	F(7,52)=8.89 p<.001	F(7,52)=8.89 p<.001	0.4 (0.12-0.83)	F(7,52)=9.06 p<.001	F(7,52)=9.06 p<.001	11.73 (6.34-17.12	F(5,62)=4.07 p<.01
	Non-Shared Narratives	102.09 (101.61-104.45)	ibd	ibd ibd	100.59 (100.05-102.24)	ibd	ibd	99.75 (97.48-102.97)	it		ibd	691 (311-1036)	ibd	4.88 (2.17-7.53)	ibd		1.38 (0.45-2.19)	ibd		28.46 (6.39-51.29	
B	Shared Narratives	105.1 (101.61-107.52)	ibd		102.09 (100.02-104.94)	ibd		101.29 (97.23-105.49)	it		ibd	876 (218-1678)	ibd	6.11 (1.13-9.57)	ibd		1.42 (0.38-3.08)	ibd		23.1 (3.8-78.12)	
C	Shared Positive Emotion	107.03 (102.09-119.76)	ibd		104.29 (101.46-112.47)	ibd		102.98 (98.74-116.74)				717 (345-1310)	ibd	7 (1.64-19.31)			1.85 (0.37-7.18)	ibd		19.4 (7.89-67.66)	
D	Shared Negative Emotion	108.76 (106.15-121.71)	ibd		105.86 (104.27-113.79)	ibd		104.72 (100.96-107.66)				1334 (330-10029)		7.52 (3.52-22.28)			1.61 (0.72-6.93)	ibd		26.14 (5.1-41.73)	
E	Squeals	120.57 (108.47-129.42)			113.76 (107.13-123.55)			115.13 (104.69-122.62)				600 (208-1159)	ibd	17.93 (1.44-25.89)			4.46 (0.48-9.96)			44.14 (20-97.43)	

Table 5.19: Julie, 7 months. Correspondence between Messages attributed by the Mother and Video Analysis.

VIDEO ANALYSIS		MESSAGE					TOTALS
		Non-Shared Narrative	Request	Shared Narrative	Shared Negative Emotion	Shared Positive Emotion	
Person-Person	Attention Seeking				3 (25)		3
	Invitation to Play			5 (17.3)		6 (24)	11
	Response to Invitation to Play					7 (28)	7
	Conflict				2 (16.6)		2
	Invitation to Stranger					2 (8)	2
Infant-Mother-Object	Converging Interest			19 (65.6)	1 (8)	4 (16)	24
	Joint Interest			1 (3.4)		3 (12)	4
	Directive					2 (8)	2
	Request for Object		4 (100)		3 (25)		7
	Split Attention			4 (13.7)		1 (4)	5
Infant-Object		7 (100)			3 (25)		10

Table 5.20: Julie, 8 months. Distribution of Messages by Contour Shape. Percentages of Contour Shapes are shown in parentheses.

MESSAGE	CONTOUR SHAPE						TOTALS
	Level	Fall	Undulating Fall	Bell	Rise	Undulating Rise	
Non-Shared Narrative			1 (25)	2 (50)		1 (25)	4
Rquest		2 (18.2)	2 (18.2)	4 (36.4)		3 (27.2)	11
Shared Narrative	1 (4)	3 (12)	8 (32)	11 (44)		2 (8)	25
Acknowledgement of Presence		4 (50)		3 (37.5)	1 (12.5.)		8
Attention Seeking		3 (25)	2 (16.7)	7 (58.3)			12
Shared Negative Emotion		1 (11.1)	5 (55.6)	2 (22.2)		1 (11.1	9
Shared Positive Emotion		3 (23.1)	2 (15.4)	6 (46.1)	2 (15.4)		13
Squeal		2 (50)		2 (50)			4

Table 5.21: Julie, 8 months. Prosodic patterns of different messages and their characteristics.

PATTERN	MESSAGE	PROSODIC FEATURES and STATISTICS																
		Peak Pitch (semitones)			Mean Pitch (semitones)			Beginning Pitch (semitones)			Final Pitch (semitones)			Duration (milliseconds)			S.D. of Pitch (semitones)	
A	Request	104.98 (98.74-110.86)	F(7,79)=7.48 p<.001		103.26 (97.26-106.31)	F(7,79)=7.48 p<.001	F(7,79)=7.48, p<.001	102.97 (97.3-106.15)	F(7,79)=8.01 p<.001		99.98 (96.72-105.22)	F(7,79)=7.78 p<.001	F(6,76)=5.59, p<.001	1126 (281-4195)			1.36 (0.43-2.77	
	Non-Shared Narratives	102.44 (99.42-104.9)	ibd	F(5,64)=8.26,p<.001 F(4,57)=5, p<.01	99.95 (98.05-102.03)		ibd F(5,64)=6.98, p<.001	98.86 (96.72-101.81)		F(6,76)=6.9, p<.001 F(3,45)=3.75, p<.05	97.55 (97.11-99.2)		ibd	807 (395-1534)			1.2 (0.75-1.58	
	Shared Narratives	103.6 (97.67-111.34)	ibd	ibd ibd	101.05 (96.87-107.36)		ibd ibd	100.21 (97.32-103.38)		ibd ibd	98.45 (96.72-104.9)		ibd	1109 (247-461)			1.28 (0.22-2.29	F(7,79)=2.94 p<.01
B	Squeals	120.63 (104.11-122.31)			116.54 (104.11-121.71)			112.61 (104.11-121.71)			114.41 (103.9-117.06)			346 (190-533)	F(7,79)=5.14 p<.001	F(5,46)=4.97, p<.001	1.57 (0.13-3.82	
	Shared Positive Emotion	109.49 (100.71-125.68)			104.52 (99.14-122.6)			103.42 (96.72-120.84)			101.81 (98.27-125.68)			395 (203-1970)		ibd	1.45 (0.52-3.41	
C	Shared Negative Emotion	110.29 (104.9-116.61)			104.93 (103.3-111.28)	ibd		106.85 (101.9-110.43)			102.38 (101.27-110.86)	ibd		2334 (678-9803)			1.66 (0.56-2.98)	
*	Attention Seeking	109.3 (104.9-115.75)			105.59 (101.08-109.29)	ibd		105.4 (100.18-112.37)			101.41 (96.72-106.3)	ibd	ibd	916 (290-2528)			2.14 (0.91-5.35)	
*	Acknowledgement of Presence	103.33 (100.45-120.57)	ibd		101.98 (99.02-115.64)	ibd		101.09 (99.31-117.06)	ibd		100.2 (97.11-113.41)	ibd		397 (250-644)	ibd	F(5,59)=3.72, p<.01 ibd	1.24 (0.83-2.56)	

Table 5.22: Julie, 8 months. Correspondence between Messages attributed by the Mother and Video Analysis.

VIDEO ANALYSIS		MESSAGE							TOTALS
		Non-Shared Narrative	Request	Shared Narrative	Shared Negative Emotion	Attention Seeking	Acknowledgement of Presence	Shared Positive Emotion	
Person-Person	Attention Seeking			2 (8)	1 (11.2)	12 (100)		4 (23.5)	19
	Response to Invitation to Play			1 (4)			2 (24)	5 (29.4)	8
	Request for Action to Self		5 (45.4)						5
	Conflict				3 (33.3)				3
	Showing Off							1 (5.9)	1
	Invitation to Stranger						1 (12.5)		1
Infant-Mother-Object	Converging Interest		2 (18.2)	18 (72)	1 (11.1)		3 (37.5)	6 (35.3)	30
	Joint Interest			3 (12)			1 (12.5)		4
	Directive				1 (11.1)			1 (5.9)	2
	Request for Object		2 (18.2)						2
	Conlfict on Object		2 (18.2)		3 (33.3)				5
Infant-Object		4 (100)		1 (4)			1 (12.5)		6

Table 5.23: Julie, 9 months. Distribution of Messages by Contour Shape. Percentages of Contour Shapes are shown in parentheses.

MESSAGE	CONTOUR SHAPE						TOTALS
	Level	Fall	Undulating Fall	Bell	Rise	Undulating Rise	
Rquest	3 (30)	3 (30)		4 (40)			10
Shared Comment		3 (42.8)		3 (42.8)		1 (14.2)	7
Shared Narrative		4 (21.1)	4 (21.1)	9 (47.4)		2 (10.4)	19
Directive		3 (33.3)		5 (55.6)	1 (11.1)		9
Acknowledgement of Presence			1 (12.5)	7 (87.5)			8
Shared Positive Emotion	1 (8.3)	4 (33.3)	4 (33.3)	2 (16.8)	1 (8.3)		12
Shared Negative Emotion			7 (50)	1 (7.2)	3 (21.4)	3 (21.4)	14

Table 5.24: Julie, 9 months. Prosodic patterns of different messages and their characteristics.

PATTERN	MESSAGE	PROSODIC FEATURES and STATISTICS																		
		Peak Pitch (semitones)			Mean Pitch (semitones)			Beginning Pitch (semitones)			Final Pitch (semitones)			Duration (milliseconds)			Pitch Range (semitones)		S.D. of Pitch (semitones)	
A	Requests	100.78 (96.92-107.03)	F(6,73)=9.76 p<.001	F(3,41)=4.28 p<.01	100.22 (96.84-106.09)	F(6,73)=10.05 p<.001	F(5,65)=7.78, p<.001 F(3.41)=4.55, p<.01	100.78 (96.92-104.65)	F(6.73)=6.33 p<.001	F(5,65)=5.12, p<.001	99.2 (96.72-105.92)	F(6,73)=5.73 p<.001	F(5,65)=4.76 p<.001	194 (152-600)	F(6.73)=9.17 p<.001	F(5,59)=8.82, p<.001 F(3,30)=4.37, p<01	1.43 (0.19-5.17)	F(6,73)=7.22 p<.001	0.45 (0.0[illegible]-1.62)	F(6,73)=6.1 p<.001
	Shared Comments	100.96 (98.15-105.69)			99.98 (97.46-104.46)		ibd ibd	99.03 (97.91-104.78)		ibd F(4,51)=5.38, p<.001	98.75 (96.72-105.22)		ibd	239 (202-505)		ibd	2.26 (1.43-3.63)		0.74 (0.[illegible]-1.13)	
B	Shared Narratives	103.6 (98.09-109.49)	ibd		101.69 (96.87-104.83)	ibd	ibd ibd	100.02 (96.92-104.24)	ibd	ibd ibd	99.2 (96.72-107.8)	ibd	ibd	880 (244-2268)			3.23 (1.37-9.31)		1.85 (0.2[illegible]-1.97)	
C	Directives	104.11 (101.27-118.74)			102.79 (100.47-117.59)			102.97 (100.18-116.84)			103.26 (99.2-118.74)			341 (236-1172)	ibd	ibd	2.99 (1.79-8.87)		1.04 (0.4[illegible]-2.53)	
D	Shared Positive Emotion	107.97 (101.71-120.02)			105.23 (100.81-118.37)			103.87 (99.08-117.98)			100.7 (96.72-118.48)			1456 (134-3346)			6.59 (1.63-20.32)		1.55 (0.4[illegible]-1.56)	
	Shared Negative Emotion	108.46 (105.22-113.8)			105.03 (102.9-107.7)			103.97 (99.25-108.37)			104.28 (102.09-108.33)			1292 (193-2486)			5.97 (2.28-12.19		1.42 (0.3[illegible]-3.83)	
*	Acknowledgement of Presence	115.44 (102.09-124.6)			112.24 (101.05-118.4)			109.98 (99.42-124.6)			107.67 (99.91-115.13)			501 (242-1294)	ibd		6.94 (1.85-14.04)		2.17 (0.6[illegible]-3.87)	

Table 5. 25: Julie, 9 months. Correspondence between Messages attributed by the Mother and Video Analysis.

VIDEO ANALYSIS		MESSAGE							TOTALS
		Shared Comment	Request	Directive	Shared Narrative	Acknowledgement of Presence	Shared Positive Emotion	Shared Negative Emotion	
Person-Person	Attention Seeking							6 (42.8)	6
	Invitation to Play				1 (5.3)		5 (41.6)		6
	Response to Invitation to Play					2 (22.2)	1 (8.3)		3
	Showing Off						1 (8.3)		1
	Conflict							2 (14.2)	2
	Invitation to Stranger					5 (55.6)			5
Infant-Mother-Object	Converging Interest	6 (85.7)			14 (73.7)	1 (11.1)	3 (25)		24
	Joint Interest	1 (14.3)	4 (40		3 (15.7)		1 (8.3)	1 (7.2)	10
	Request for Object		2 (20)						2
	Directive			8 (88.9)	1 (5.3)		1 (8.3)		10
	Conflict on Object		4 (40)					4 (28.6)	8
Infant-Object				1 (11.1)		1 (11.1)		1 (7.2)	3

Table 5.26: Julie, 10 months. Distribution of Messages by Contour Shape. Percentages of Contour Shapes are shown in parentheses.

MESSAGE	CONTOUR SHAPE						TOTALS
	Level	Fall	Undulating Fall	Bell	Rise	Undulating Rise	
Rquest	3 (50)	2 (33.3)		1 (16.7)			6
Shared Interest		3 (60)		2 (40)			5
Directive	2 (10)	13 (65)		4 (20)	1 (5)		20
Response		3 (60)		2 (40)			5
Shared Comment				4 (100)			4
Shared Positive Emotion		2 (25)		5 (62.5)	1 (12.5)		8
Shared Narrative		1 (9.1)	7 (63.6)	3 (26.6)			11
Shared Negative Emotion			7 (70)	2 (20)	1 (10)		10

Table 5.27: Julie, 10 months. Prosodic patterns of different messages and their characteristics.

PATTERN	MESSAGE	PROSODIC PATTERNS and STATISTICS												
		Peak Pitch (semitones)		Mean Pitch (semitones)		Beginning Pitch (semitones)		Duration (milliseconds)			Pitch Range (semitones)		S.D. of Pitch (semitones)	
A	Requests	101.95 (99.25-105.45)	F(7,62)=4.62 p<.001	101.18 (98.73-105.06)	F(7,62)=4.62 p<.001	100.55 (98.8-105.45)		167 (111-313)	F(7,62)=36.41 p<.001	F(7,62)=36.41, p<.001 F(5,43)=6, p<.001	1.69 (0.67-2.91)	F(7,62)=9 p<.001	0.55 (0.2-0.89)	F(7,62)=5.33 p<.0
	Shared Interest	103.29 (98.86-106.26)		101.35 (98.02-102.67)		101.41 (98.15-106.26)		327 (236-580)		ibd	5.49 (1.63-6.39)		1.4 (0.39-3.41)	
	Directives	101.09 (97.42-108.47)		99.52 (96.92-105.67)		100.06 (96.72-108.47)	F(7,62)=3.68 p<.01	249 (146-543)		ibd	1.93 (0.34-6.86)	ibd	0.74 (0.13-2.42)	ibd
	Responses	101.9 (99.8-103.98)		99.64 (98.65-103.14)		101.51 (97.42-103.6)		301 (171-326)		ibd	2.38 (1.13-4.22)	ibd	0.77 (0.31-1.51)	ibd
	Shared Comments	101.09 (100.18-102.38)		100.31 (99.64-101.45)		98.98 (98.27-100.5)	ibd	244 (209-273)		ibd	1.97 (1.9-2.3)	ibd	0.54 (0.46-0.73)	ibd
B	Shared Narratives	104.9 (99.91-117.06)		102.75 (97.9-109.31)		102 (98.09-111.51)		1608 (500-4170)			5.88 (2.9-17.14)		1.56 (0.85-5.51)	
C	Shared Negative Emotion	109.12 (106.77-109.92)		105.57 (103.06-107.17)		105.64 (102.79-108.63)		3453 (612-9570)			6.07 (2.53-8.87)		1.52 (0.58-2.62)	
*	Shared Positive Emotion	104.79 (101.81-107.8)		103.22 (99.71-104.85)		103.51 (99.31-106.56)		506 (224-593)	ibd		4.31 (1.74-5.86)		1.32 (0.56-2.04)	

Table 5.28: Julie, 10 months. Correspondence between Messages attributed by the Mother and Video Analysis.

VIDEO ANALYSIS		MESSAGE								TOTALS
		Request	Response	Shared Comment	Directive	Shared Interest	Shared Narrative	Shared Positive Emotion	Shared Negative Emotion	
Pesron-Person	Attention Seeking								8 (80)	8
	Response to Invitation to Play							4 (50)		4
	Showing Off		2 (20)							2
	Conflict								2 (20)	2
Mother-Infant-Object	Converging Interest	1 (16.6)	2 (40)	2 (40)		1 (20)	11 (100)			6
	Joint Interest		2 (40)	3 (60)		3 (60)		2 (25)		10
	Directive				20 (100)			2 (25)		22
	Request for Object	1 (16.6)								1
	Conflict on Object	4 (66.8)								4
	Split Interest					1 (20)				1

Table 5.29: Julie, 11 months. Distribution of Messages by Contour Shape. Percentages of Contour Shapes are shown in parentheses.

MESSAGE	CONTOUR SHAPE						TOTALS
	Level	Fall	Undulating Fall	Bell	Rise	Undulating Rise	
Non-Shared Narrative		2 (16.7)	5 (41.6)	2 (16.7)		3 (25	12
Shared Interest	4 (36.4)	5 (45.5)		2 (18.1)			11
Directive	3 (60)	2 (40)					5
Response	4 (57.1)	2 (28.6)		1 (14.3)			7
Shared Comment	2 (50)	1 (25)		1 (25)			4
Attention Seeking	2 (28.6)	1 (14.2)	2 (28.6)	2 (28.6)			7
Shared Narrative			8 (66.7)	1 (8.3)		3 (25)	12
Shared Positive Emotion	2 (9.1)	7 (31.8)		13 (59.1)			22

Table 5.30: Julie, 11 months. Prosodic patterns of different messages and their characteristics.

PATTERN	MESSAGE	PROSODIC FEATURES and STATISTICS																		
		Peak Pitch (semitones)			Mean Pitch (semitones)			Beginning Pitch (semitones)		Final Pitch (semitones)		Duration (milliseconds)		Pitch Range (semitones)			S. D. of Pitch (semitones)		Rate of Pitch Change (milliseconds/sec)	
A	Directives	100.29 (97.05-100.96)	F(7,73)=7.01 p<.001	F(5,40)=4.69 p<.01	99.35 (96.82-99.5)	F(7,73)=8.03 p<.001		99.92 (96.72-100.96)	F(7,73)=6.27 p<.001	97.05 (96.79-99.08)	F(7,73)=5.62 p<.001	220 (172-228)	F(7,73)=25.22 p<.001	2.8 (0.11-4.1)	F(7,73)=8.39 p<.001		0.86 (0.05-1.48)		31 (20.79-41.0[illegible]	F(5,41)=4.1 p<.005
	Shared Comments	97.85 (96.92-100.96)		ibd	97.41 (96.76-99.44)		F(6,51)=2.95 p<.01	97.5 (96.92-98.74)		97.05 (96.72-97.73)		200 (150-219)		0.8 (0.19-3.22)	ibd		0.3 (0.08-0.96)	F(7,73)=7 p<.001	32.65 (16.09-49.2[illegible]	
	Responses	100.71 (97.05-102.5)			99.04 (96.76-100.96)			99.6 (96.72-101.06)		97.11 (96.72-101.06)		193 (95-299)		0.6 (0.2-3.78)	ibd	F(7,73)=8.39 p<.001	0.21 (0.08-1.42)	ibd	20.95 (13.18-25.1[illegible]	
	Shared Interest	101.51 (98.03-105.69)			100.01 (97.48-103.72)			99.97 (96.72-105.1)		98.04 (96.72-101.51)		259 (191-716)		1.73 (0.95-7.06)			0.54 (0.36-2.27)		17.89 (8.8-31.63	
B	Shared Narratives	104.08 (99.58-116.41)			100.6 (97.84-106.4)	ibd		100.26 (96.72-103.6)	ibd	97.56 (96.72-102.66)	ibd	2142 (591-3935)		5.58 (2.86-19.68)			1.55 (0.77-5.75)		15.39 (13.21-25.8	ibd
C	Shared Positive Emotion	105.16 (102.75-116.17)			103.62 (99.42-113.91)			103.44 (99.08-113.59)		101.61 (97.24-113.68)		284 (160-1218)	ibd	4.79 (1.12-10.93)			1.36 (0.36-3.17)		28.51 (4.82-45.35	
*	Non-Shared Narratives	102.28 (100.18-110.86)			99.96 (98.44-103.66)	ibd		100.73 (96.72-109.5)		98.83 (96.72-103.6)	ibd	2709 (275-5329)		5.15 (2.68-10.41)			1.37 (0.73-1.89)		13.16 (3.69-30.31	ibd
*	Attention Seeking	102.47 (100.71-107.8)			101.29 (98.49-105.13)			100.6 (96.72-102.09)	ibd	100.86 (96.72-103.68)		330 (186-1680)	ibd	3.98 (0.5-9.52)			1.14 (0.17-2.31)		58.05 (49.46-97.4	

Table 5.31: Julie, 11 months. Correspondence between Messages attributed by the Mother and Video Analysis.

VIDEO ANALYSIS		MESSAGE								TOTALS
		Non-Shared Narrative	Shared Narrative	Shared Comment	Response	Directive	Shared Interest	Attention Seeking	Shared Positive Emotion	
Person-Person	Attention Seeking							7 (100)		7
	Teasing				1 (14.3)					1
	Invitation to Play								4 (18.2)	4
	Showing Off		1 (8.3)							1
	Response to Invitation to Play		2 (16.7)						13 (59.1)	15
Mother-Infant-Object	Converging Interest		9 (75)	3 (75)	1 (14.3)		3 (27.3)			16
	Joint Interest			1 (25)	4 (57.1)		8 (72.7)		3 (13.6)	16
	Directive				1 (14.3)	5 (100)			2 (9.1)	8
Infant-Object		12 (100)								12

APPENDIX III

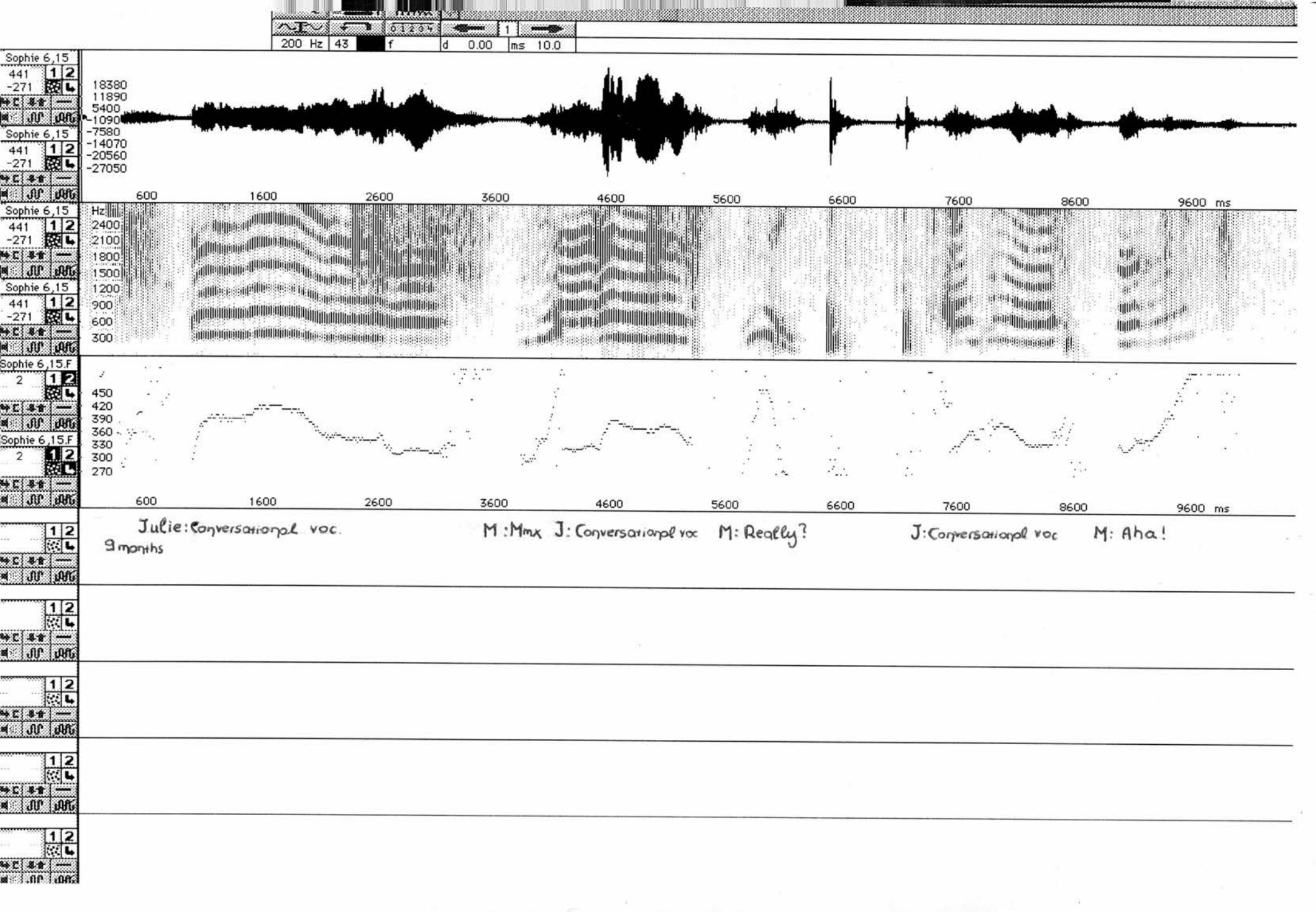
ILLUSTRATIONS

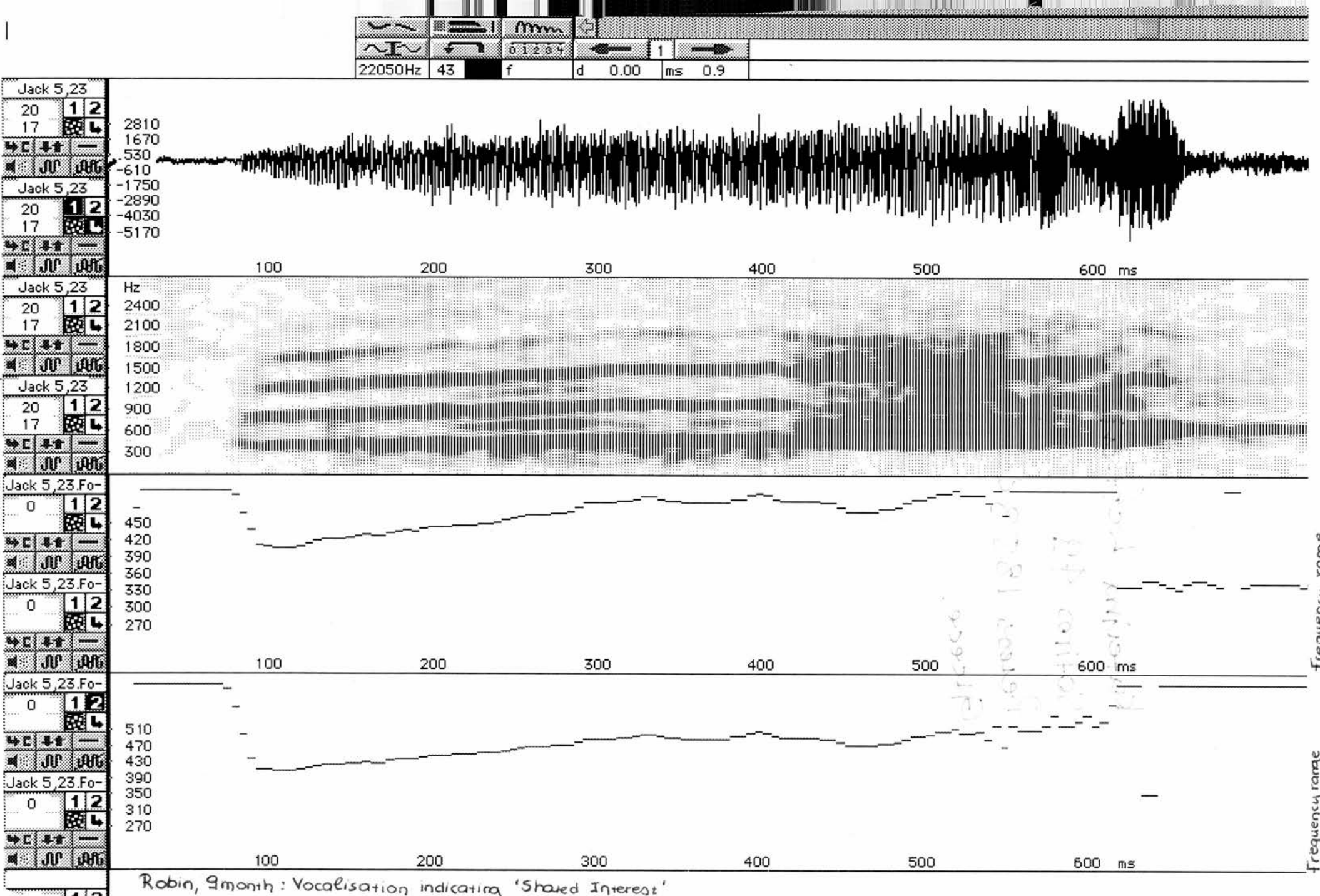
Illustration 1: Robin, 7 months. Taking a toy the mother is offering.

Illustration 2: Julie, 8 months. Acknowledging the presence of the Stranger.

Illustration 3: Robin, 10 months. Alternating gaze between the mother and the toy she is showing.

Illustration 4: Julie, 10 months. Showing a toy to the mother.


200 Hz
43
f
d 0.00
ms 10.0
Sophie 6,15
441
-271
18380
11890
5400
-1090
-7580
-14070
-20560
-27050
600
1600
2600
3600
4600
5600
6600
7600
8600
9600 ms
Hz
2400
2100
1800
1500
1200
900
600
300
Sophie 6,15.F
2
450
420
390
360
330
300
270
Julie: Conversational voc.
9 months
M: Mmx J: Conversational voc
M: Really?
J: Conversational voc
M: Aha!



Robin, 9month : Vocalisation indicating 'Shared Interest'